

```

!!SEQUENCE List 1.0
! FINDPATTERNS on geneseq: * allowing 0 mismatches
1 <(X){1,200}(L,I,V,M,A,P)(X,T,S)(L,I,V,M,A,P)(X)(L,I,V,M,A,F,Y,W)(C,T,S)(R,K,H

GENESEQ:P90009 CK: 7203 len: 102 finds: 1 | Mouse gamma-2b chain New Immun
GENESEQ:P91884 CK: 718 len: 48 finds: 1 | Antigenic Epstein-Barr virus F
GENESEQ:P94157 CK: 2829 len: 246 finds: 1 | Acetyl-CoA reductase. Construc
GENESEQ:P90460 CK: 560 len: 288 finds: 1 | Alpha-factor profilinase from
GENESEQ:P81996 CK: 4995 len: 25 finds: 1 | Binding peptide 2 capable of h
GENESEQ:P70477 CK: 9828 len: 194 finds: 1 | Sequence of human respiratory
GENESEQ:P70785 CK: 9828 len: 194 finds: 1 | Sequence encoding human respi
GENESEQ:P24296 CK: 3600 len: 384 finds: 1 | Regulatory protein Vans Invol
GENESEQ:R10995 CK: 1620 len: 398 finds: 1 | Bovine RSV strain A 51908 M2 F
GENESEQ:R24190 CK: 5872 len: 186 finds: 1 | Sequence of rhinovirus HRV2 v
GENESEQ:P81092 CK: 3060 len: 95 finds: 1 | Sequence of rhinovirus HRV2 v
GENESEQ:P81093 CK: 5281 len: 323 finds: 1 | Acetoacetyl CoA reductase enzy
GENESEQ:R10974 CK: 3836 len: 246 finds: 1 | Luffa cylindrica bioactive pro
GENESEQ:R12468 CK: 7673 len: 248 finds: 1 | Murine Cytotoxic Cell Protease
GENESEQ:R13252 CK: 528 len: 247 finds: 1 | Mouse kidney cell Band 3-like
GENESEQ:P60645 CK: 2381 len: 292 finds: 1 | HCV-T (6110-6557) encoded epit
GENESEQ:R12427 CK: 3078 len: 149 finds: 1 | VP6 inner capsid epitope, 84 T
GENESEQ:R20004 CK: 4995 len: 25 finds: 1 | NADH dehydrogenase subunit 2.
GENESEQ:R21409 CK: 617 len: 482 finds: 1 | Protein transcribed from the m
GENESEQ:R23004 CK: 7059 len: 297 finds: 1 | blac31 mutation (5). Novel DNA
GENESEQ:R22695 CK: 4919 len: 286 finds: 1 | NANBH virus antigen. Antigenic
GENESEQ:R25601 CK: 5700 len: 85 finds: 1 | HCV polypeptide 16. Hepatitis
GENESEQ:R25869 CK: 3783 len: 239 finds: 1 | HCV NS4-NS5 peptide N29-1. New
GENESEQ:R25891 CK: 4644 len: 253 finds: 1 | HCV NS4-NS5 peptide N29-2. New
GENESEQ:R25303 CK: 9668 len: 194 finds: 1 | HCV NS4-NS5 peptide N29-3. New
GENESEQ:R29874 CK: 9846 len: 171 finds: 1 | HCV NS4-NS5 peptide N29-3. New
GENESEQ:R29879 CK: 8945 len: 285 finds: 1 | HCV NS4-NS5 peptide N29-3. New
GENESEQ:R29895 CK: 8862 len: 171 finds: 1 | HCV NS4-NS5 peptide N29-3. New
GENESEQ:R29896 CK: 8862 len: 171 finds: 1 | HCV NS4-NS5 peptide N29-3. New
GENESEQ:R29897 CK: 2886 len: 167 finds: 1 | Prod. of the luffin-g gene. Ne
GENESEQ:R29909 CK: 429 len: 277 finds: 1 | Prod. of the luffin-g gene. Ne
GENESEQ:R29910 CK: 9157 len: 278 finds: 1 | Homologous to chicken nov gene
GENESEQ:R31608 CK: 7012 len: 205 finds: 1 | Polyptide coded by Korean HO
GENESEQ:R30618 CK: 7998 len: 106 finds: 1 |

```

```

GENESEQ:R32192 CK: 2829 len: 246 finds: 1 | Sequence encoded by the ace
GENESEQ:R37294 CK: 7673 len: 248 finds: 1 | Plant type I RIP luffin. An
GENESEQ:R39259 CK: 8652 len: 391 finds: 1 | Human somatostatin receptor
GENESEQ:R39260 CK: 8110 len: 391 finds: 1 | Murine somatostatin recepto
GENESEQ:R42317 CK: 3489 len: 345 finds: 1 | EBV VCA-P40 from ORF BBRF1.
GENESEQ:R52667 CK: 5723 len: 303 finds: 1 | Equine herpesvirus US2 gene
GENESEQ:R49296 CK: 4901 len: 25 finds: 1 | Invariant chain (Ii) positi
GENESEQ:R49297 CK: 3126 len: 24 finds: 1 | Invariant chain (Ii) positi
GENESEQ:R49299 CK: 1278 len: 23 finds: 1 | Invariant chain (Ii) positi
GENESEQ:R49438 CK: 4300 len: 49 finds: 1 | Maligne HLA-Dalpha chain/
GENESEQ:R49589 CK: 4300 len: 49 finds: 1 | Sequence of HLA-DR alpha ch
GENESEQ:R50081 CK: 3747 len: 352 finds: 1 | NANBH virus antigenic fragm
GENESEQ:R50082 CK: 108 len: 215 finds: 1 | NANBH virus antigenic fragm
GENESEQ:R53251 CK: 2035 len: 56 finds: 1 | Consensus sequence of signa
GENESEQ:R53646 CK: 4677 len: 380 finds: 1 | c-fos gene product. New ant
GENESEQ:R53748 CK: 2863 len: 355 finds: 1 | Seven transmembrane recepto
GENESEQ:R58556 CK: 4528 len: 175 finds: 1 | Heparin-binding secretory t
GENESEQ:R63592 CK: 2807 len: 79 finds: 1 | MAP-kinase-phosphatase CL10
GENESEQ:R63602 CK: 5661 len: 367 finds: 1 | MAP-kinase-phosphatase CL10
GENESEQ:R68635 CK: 2665 len: 124 finds: 1 | Amylase inhibitor protein 0
GENESEQ:R68636 CK: 3320 len: 124 finds: 1 | Amylase inhibitor protein 0
GENESEQ:R68687 CK: 9708 len: 246 finds: 1 | Nicotinicamide adenine dinuc
GENESEQ:R63443 CK: 8722 len: 120 finds: 1 | Amino terminal sequence of
GENESEQ:R63444 CK: 7811 len: 120 finds: 1 | Amino terminal sequence of
GENESEQ:R62507 CK: 2303 len: 313 finds: 1 | Galactosyl transferase 3' c
GENESEQ:R66355 CK: 1046 len: 216 finds: 1 | IL protein. Identification
GENESEQ:R63159 CK: 2293 len: 376 finds: 1 | Mouse growth differentiatio
GENESEQ:R63160 CK: 1814 len: 375 finds: 1 | Human growth differentiatio
GENESEQ:R66665 CK: 4995 len: 25 finds: 1 | Binding peptide B (84 TS) I
GENESEQ:R66680 CK: 5136 len: 25 finds: 1 | Peptide MONOSER. Immunogeni
GENESEQ:R67596 CK: 5465 len: 331 finds: 1 | A. aculeatus pectin methyl
GENESEQ:R71324 CK: 7564 len: 329 finds: 1 | Acetyl-CoA-reductase. Trans
GENESEQ:R70100 CK: 4652 len: 274 finds: 1 | Lettuce infectious yellows
GENESEQ:R72715 CK: 516 len: 359 finds: 1 | hsc gene product of Salmon
GENESEQ:R74180 CK: 8163 len: 248 finds: 1 | Type I ribosome-inactivatin

```

GENESOP:R84103	ck: 5723	len: 303	finds: 1	I Equine herpesvirus (EHV) unid	GENESOP:W20259	ck: 5141	len: 108	finds: 1	I H. pylori secreted or perip
GENESOP:R81323	ck: 9504	len: 123	finds: 1	I Humanized VLA-4 antibody 21.6	GENESOP:W20324	ck: 6372	len: 102	finds: 1	I H. pylori cell envelope tra
GENESOP:R81327	ck: 2629	len: 140	finds: 1	I Mouse VLA-4 antibody 21.6 11g	GENESOP:W20696	ck: 7893	len: 121	finds: 1	I H. pylori secreted or perip
GENESOP:R81330	ck: 8013	len: 123	finds: 1	I Mouse anti-VLA-4 antibody 21.6	GENESOP:W20139	ck: 1367	len: 53	finds: 1	I H. pylori cytoplasmic prote
GENESOP:R81333	ck: 9019	len: 142	finds: 1	I Human VLA-4 reshaped antibody	GENESOP:W09433	ck: 8653	len: 328	finds: 1	I Human placenta purinergic P
GENESOP:R75202	ck: 824	len: 225	finds: 1	I Tyrosine phosphatase MPP-delta	GENESOP:W21704	ck: 7673	len: 248	finds: 1	I Luffin-A. Inactive precursor
GENESOP:R91274	ck: 1456	len: 62	finds: 1	I Elmeria gametocyte antigen. P	GENESOP:W21709	ck: 7681	len: 250	finds: 1	I Luffin-B. Inactive precursor
GENESOP:R94989	ck: 5431	len: 354	finds: 1	I NSK2 receptor intracellular do	GENESOP:W10031	ck: 8373	len: 74	finds: 1	I Protein encoded by clone G-
GENESOP:R94990	ck: 5928	len: 282	finds: 1	I NSK2 receptor tyrosine-kinase	GENESOP:W13647	ck: 5362	len: 30	finds: 1	I Invariant chain region whic
GENESOP:R89896	ck: 6122	len: 423	finds: 1	I Rat kynurenine aminotransferas	GENESOP:W25141	ck: 7673	len: 248	finds: 1	I Luffin-A (a ribosome inhibi
GENESOP:R89897	ck: 206	len: 437	finds: 1	I Rat kynurenine aminotransferas	GENESOP:W25146	ck: 7681	len: 250	finds: 1	I Luffin-B (a ribosome inhibi
GENESOP:R89898	ck: 7276	len: 457	finds: 1	I Rat kynurenine aminotransferas	GENESOP:W22410	ck: 2629	len: 140	finds: 1	I Alpha-4 integrin mouse Mab
GENESOP:R77433	ck: 2940	len: 160	finds: 1	I Antigen of Rochalimaea henselae	GENESOP:W22413	ck: 9557	len: 123	finds: 1	I Humanised alpha-4 integrin
GENESOP:R95054	ck: 9286	len: 342	finds: 1	I TGF-alpha-DETA-DGAL4 multidomain p	GENESOP:W22428	ck: 9019	len: 142	finds: 1	I Humanised alpha-4 integrin
GENESOP:R95055	ck: 7433	len: 421	finds: 1	I IL-2-DETA-DGAL4 multidomain p	GENESOP:W26413	ck: 3737	len: 121	finds: 1	I Swinepox virus HindIII C en
GENESOP:R91225	ck: 7811	len: 338	finds: 1	I Human placenta G-protein coupl	GENESOP:W17971	ck: 6033	len: 47	finds: 1	I RAC protein kinase C-termin
GENESOP:R99695	ck: 4995	len: 25	finds: 1	I VP6 binding peptide B. Encapsu	GENESOP:W34668	ck: 9097	len: 438	finds: 1	I Arabidopsis thaliana Rec-A
GENESOP:R99501	ck: 2030	len: 244	finds: 1	I Nitrlase regulatory factor 24	GENESOP:W24491	ck: 1423	len: 122	finds: 1	I Novel amylase inhibitor pro
GENESOP:R99576	ck: 2766	len: 83	finds: 1	I Waap venom Brhnx-1 subunit (b)	GENESOP:W24492	ck: 2685	len: 124	finds: 1	I Novel amylase inhibitor pro
GENESOP:W00624	ck: 8391	len: 358	finds: 1	I Sacl endonuclease. DNA encodin	GENESOP:W24493	ck: 3320	len: 124	finds: 1	I Novel amylase inhibitor pro
GENESOP:R94579	ck: 2400	len: 488	finds: 1	I Chlamydia pneumoniae polypepti	GENESOP:W29523	ck: 2685	len: 124	finds: 1	I Wheat amylase inhibitor 0.2
GENESOP:R94580	ck: 3773	len: 271	finds: 1	I C. pneumoniae polypeptide anti	GENESOP:W29524	ck: 1423	len: 122	finds: 1	I Wheat amylase inhibitor 0.2
GENESOP:R94586	ck: 9728	len: 259	finds: 1	I C. pneumoniae polypeptide anti	GENESOP:W29525	ck: 3320	len: 124	finds: 1	I Wheat amylase inhibitor 0.1
GENESOP:R99316	ck: 4323	len: 458	finds: 1	I Human SH-PTP1 variant derived	GENESOP:R63906	ck: 7673	len: 248	finds: 1	I Type I ribosome-inactivatin
GENESOP:R83015	ck: 5439	len: 371	finds: 1	I Human thyroid transcription fa	GENESOP:W45093	ck: 9381	len: 176	finds: 1	I Residues 217-392 of human t
GENESOP:W00725	ck: 1679	len: 207	finds: 1	I Vascular endothelial growth fa	GENESOP:W38190	ck: 1840	len: 107	finds: 1	I Soybean SCARECROW SRP1 pro
GENESOP:R99253	ck: 4145	len: 374	finds: 1	I Cytoplasmic antiprotease-2 p	GENESOP:W34214	ck: 5054	len: 271	finds: 1	I Streptomyces ketoreductase
GENESOP:W01743	ck: 9728	len: 259	finds: 1	I C. pneumoniae 53 kDa antigen.	GENESOP:W41747	ck: 3783	len: 239	finds: 1	I Hepatitis C virus antigen.
GENESOP:W04831	ck: 1679	len: 207	finds: 1	I Vascular endothelial growth fa	GENESOP:W47606	ck: 9828	len: 194	finds: 1	I HRSV protein 22K. Productio
GENESOP:W11481	ck: 5195	len: 205	finds: 1	I D. immitis mature venom allerg	GENESOP:W50006	ck: 5567	len: 365	finds: 1	I Human papillomavirus-18 E2
GENESOP:W11479	ck: 9389	len: 221	finds: 1	I D. immitis venom allergen anti	GENESOP:W55590	ck: 428	len: 307	finds: 1	I H. pylori ORF 09cpl0224_429
GENESOP:W13772	ck: 2018	len: 244	finds: 1	I Rhodococcus erythropolis SK92-	GENESOP:W55815	ck: 5054	len: 271	finds: 1	I Streptomyces roseofulvus fr
GENESOP:W06943	ck: 9316	len: 140	finds: 1	I Cagl locus product 14. Helicob	GENESOP:W56137	ck: 7408	len: 186	finds: 1	I Open reading frame 2, peptid
GENESOP:W06935	ck: 8017	len: 136	finds: 1	I Cagl locus product 6. Helicoba	GENESOP:W51011	ck: 7216	len: 277	finds: 1	I Human liver carbonyl reduct
GENESOP:W20571	ck: 6963	len: 114	finds: 1	I H. pylori secreted or periplas	GENESOP:W28052	ck: 2273	len: 269	finds: 1	I Amino acid sequence of phos
GENESOP:W20886	ck: 2872	len: 154	finds: 1	I H. pylori secreted or periplas	GENESOP:W37816	ck: 3813	len: 317	finds: 1	I Human secreted apoptosis-re
GENESOP:W20776	ck: 6185	len: 276	finds: 1	I H. pylori flagella-associated					

GENESEOP:W59024	ck: 9911	len: 71	finds: 1	Enterocin-900. Bacterial growth
GENESEOP:W57330	ck: 4443	len: 339	finds: 1	Glycerol-3-phosphate dehydrog
GENESEOP:W62641	ck: 7060	len: 137	finds: 1	Fla serine protease inhibitor
GENESEOP:W55052	ck: 4306	len: 118	finds: 1	Sunflower antifungal protein M
GENESEOP:W48722	ck: 2863	len: 355	finds: 1	Human V28 seven transmembrane
GENESEOP:W54728	ck: 3126	len: 24	finds: 1	Peptide from IL p80-103. Incre
GENESEOP:W60258	ck: 4443	len: 339	finds: 1	Klebsiella pneumoniae glycerol
GENESEOP:W60267	ck: 843	len: 387	finds: 1	Klebsiella pneumoniae DHAT pro
GENESEOP:W60764	ck: 7422	len: 260	finds: 1	Rainbow trout interleukin 1 be
GENESEOP:W57889	ck: 5290	len: 271	finds: 1	Corn raffinose synthetase. New
GENESEOP:W44905	ck: 1135	len: 39	finds: 1	"polypyrrolone beta-turn helix"
GENESEOP:W44906	ck: 6449	len: 70	finds: 1	"polypyrrolone beta-turn helix"
GENESEOP:W44907	ck: 3740	len: 127	finds: 1	"polypyrrolone beta-turn helix"
GENESEOP:W44908	ck: 1205	len: 79	finds: 1	"polypyrrolone beta-turn helix"
GENESEOP:W44909	ck: 4699	len: 110	finds: 1	"polypyrrolone beta-turn helix"
GENESEOP:W44910	ck: 1990	len: 167	finds: 1	"polypyrrolone beta-turn helix"
GENESEOP:W44879	ck: 9655	len: 104	finds: 1	"polypyrrolone beta-turn helix"
GENESEOP:W44882	ck: 8918	len: 61	finds: 1	"polypyrrolone beta-turn helix"
GENESEOP:W62612	ck: 8912	len: 369	finds: 1	Human glutamate-binding protei
GENESEOP:W30679	ck: 843	len: 387	finds: 1	1,3-Propanediol oxidoreductase
GENESEOP:W62623	ck: 3029	len: 263	finds: 1	For-II protein sequence. Nucle
GENESEOP:W62617	ck: 5610	len: 212	finds: 1	Rattus norvegicus SOCS1 protei
GENESEOP:W62622	ck: 4000	len: 130	finds: 1	Mus musculus SOCS8 protein. Su
GENESEOP:W62623	ck: 6307	len: 207	finds: 1	Homo sapiens SOCS11 protein. S
GENESEOP:W62624	ck: 8778	len: 134	finds: 1	Mus musculus SOCS13 protein. S
GENESEOP:W62613	ck: 4527	len: 212	finds: 1	Mus musculus SOCS1 protein. Su
GENESEOP:W62614	ck: 1060	len: 198	finds: 1	Mus musculus SOCS2 protein. Su
GENESEOP:W62615	ck: 9828	len: 225	finds: 1	Mus musculus SOCS3 protein. Su
GENESEOP:W62616	ck: 758	len: 211	finds: 1	Homo sapiens SOCS1 protein. Su
GENESEOP:W61619	ck: 1162	len: 250	finds: 1	Clone HTPEF86 of TM4SF superfa
GENESEOP:W65460	ck: 1814	len: 375	finds: 1	Human growth differentiation f
GENESEOP:W64373	ck: 3575	len: 120	finds: 1	Mycobacterium tuberculosis ant
GENESEOP:W38580	ck: 5404	len: 174	finds: 1	Streptococcus pneumoniae proce
GENESEOP:W70962	ck: 4556	len: 212	finds: 1	A STAT function regulatory pro
GENESEOP:W77666	ck: 5825	len: 100	finds: 1	Staphylococcus aureus protein
GENESEOP:W38716	ck: 3773	len: 161	finds: 1	S. pneumoniae aspartyl tRNA sy
GENESEOP:W79302	ck: 8790	len: 67	finds: 1	A Staphylococcus aureus protei
GENESEOP:W69889	ck: 1496	len: 376	finds: 1	Rat growth differentiation
GENESEOP:W69881	ck: 1805	len: 375	finds: 1	Pig growth differentiation
GENESEOP:W69892	ck: 1229	len: 375	finds: 1	Ovine growth differentiation
GENESEOP:W30689	ck: 2293	len: 376	finds: 1	Murine growth differentiat
GENESEOP:W69885	ck: 1814	len: 375	finds: 1	Human growth differentiat
GENESEOP:W69886	ck: 1463	len: 375	finds: 1	Baboon growth differentiat
GENESEOP:W69887	ck: 9305	len: 375	finds: 1	Bovine growth differentiat
GENESEOP:W80799	ck: 2018	len: 244	finds: 1	Rhodococcus nitrite hydrata
GENESEOP:W80624	ck: 4504	len: 189	finds: 1	S. pneumoniae GRP binding p
GENESEOP:W68494	ck: 9982	len: 368	finds: 1	E2 papillomavirus protein e
GENESEOP:W82504	ck: 447	len: 225	finds: 1	Human EPRG1 protein #1. New
GENESEOP:W82505	ck: 9756	len: 157	finds: 1	Human EPRG1 protein #2. New
GENESEOP:W81740	ck: 3575	len: 120	finds: 1	M. tuberculosis immunogenic
GENESEOP:W72902	ck: 5027	len: 226	finds: 1	Mycobacterium tuberculosis
GENESEOP:W80495	ck: 1679	len: 207	finds: 1	Human vascular endothelial
GENESEOP:W85306	ck: 3126	len: 24	finds: 1	Helper T-cell class II pept
GENESEOP:W86202	ck: 9273	len: 206	finds: 1	Human VEGF-related factor (
GENESEOP:W68541	ck: 8848	len: 420	finds: 1	Amino acid sequence of the
GENESEOP:W86235	ck: 8227	len: 185	finds: 1	Human VRF (VGF-related fac
GENESEOP:W86214	ck: 5797	len: 178	finds: 1	Human VRF-2 truncated fragm
GENESEOP:W86215	ck: 4363	len: 173	finds: 1	Human VRF-2 truncated fragm
GENESEOP:W86216	ck: 2203	len: 168	finds: 1	Human VRF-2 truncated fragm
GENESEOP:W86217	ck: 3123	len: 163	finds: 1	Human VRF-2 truncated fragm
GENESEOP:W85119	ck: 5663	len: 446	finds: 1	A delta-5 desaturase enzyme
GENESEOP:W73507	ck: 3945	len: 317	finds: 1	Human ATG-1709 protein. New
GENESEOP:W90011	ck: 9038	len: 343	finds: 1	Expressed antigen for clust
GENESEOP:W89910	ck: 8829	len: 343	finds: 1	Antigen 1 from cluster 35a.
GENESEOP:W85590	ck: 2306	len: 354	finds: 1	Polypeptide with 98% homolo
GENESEOP:W85591	ck: 8607	len: 351	finds: 1	Polypeptide with 95% homolo
GENESEOP:W89100	ck: 1342	len: 36	finds: 1	Polypeptide fragment encode
GENESEOP:W88568	ck: 9823	len: 221	finds: 1	Secreted protein encoded by
GENESEOP:W83391	ck: 9212	len: 161	finds: 1	Caenorhabditis elegans sym
GENESEOP:W95506	ck: 5663	len: 446	finds: 1	Mortierella alpina delta 5
GENESEOP:W98862	ck: 4789	len: 304	finds: 1	H. pylori GHP0 1670 protein
GENESEOP:W98755	ck: 3203	len: 102	finds: 1	H. pylori GHP0 1075 protein
GENESEOP:W98226	ck: 2117	len: 115	finds: 1	H. pylori GHP0 1099 protein

GENESEQP:W98330	ck: 7734	len: 336	finds: 1	! Mannosyl transferase involved
GENESEQP:Y01440	ck: 9731	len: 64	finds: 1	! Secreted protein encoded by gene
GENESEQP:W99185	ck: 9793	len: 314	finds: 1	! Rhodococcus corallina ohp oper
GENESEQP:W97722	ck: 1929	len: 302	finds: 1	! Staphylococcus aureus mutant F
GENESEQP:Y11009	ck: 2840	len: 142	finds: 1	! H. pylori ORF hp6p22217_2356440
GENESEQP:Y11062	ck: 4555	len: 116	finds: 1	! H. pylori ORF 06cp30603_1074440
GENESEQP:W97887	ck: 1814	len: 375	finds: 1	! Human myostatin. Increasing m
GENESEQP:W97884	ck: 8371	len: 375	finds: 1	! Bovine myostatin. Increasing m
GENESEQP:W97885	ck: 1462	len: 286	finds: 1	! Bovine myostatin (mutant form)
GENESEQP:W97886	ck: 2420	len: 376	finds: 1	! Murine myostatin. Increasing m
GENESEQP:W97840	ck: 7697	len: 207	finds: 1	! Winged bean chymotrypsin inhibi
GENESEQP:Y03794	ck: 7277	len: 144	finds: 1	! S. aureus polypeptide. New esse
GENESEQP:Y03782	ck: 881	len: 135	finds: 1	! S. aureus polypeptide. New esse
GENESEQP:Y02934	ck: 2639	len: 53	finds: 1	! Fragment of human secreted pro
GENESEQP:Y11970	ck: 1817	len: 113	finds: 1	! Human 5' ESR secreted protein
GENESEQP:W96262	ck: 3324	len: 423	finds: 1	! Bcr-3a polypeptide. New polype
GENESEQP:Y05204	ck: 3125	len: 24	finds: 1	! Human CLIP immunomodulatory p
GENESEQP:Y11993	ck: 4119	len: 102	finds: 1	! Human 5' ESR secreted protein
GENESEQP:Y05545	ck: 5919	len: 243	finds: 1	! Wheat Type III glutathione tr
GENESEQP:Y07041	ck: 1679	len: 207	finds: 1	! Breast cancer associated anti
GENESEQP:Y07007	ck: 4692	len: 177	finds: 1	! Breast cancer associated anti
GENESEQP:W93953	ck: 7759	len: 352	finds: 1	! Human regulatory molecule HRM
GENESEQP:Y04959	ck: 8454	len: 273	finds: 1	! Mycobacterium species protein
GENESEQP:Y04960	ck: 8320	len: 280	finds: 1	! Mycobacterium species protein
GENESEQP:Y04901	ck: 6286	len: 72	finds: 1	! Mycobacterium species protein
GENESEQP:Y04743	ck: 5429	len: 110	finds: 1	! Mycobacterium species protein
GENESEQP:Y07239	ck: 2144	len: 141	finds: 1	! Fragment of human placental s
GENESEQP:Y07240	ck: 2614	len: 141	finds: 1	! Fragment of human hepatic STAR
GENESEQP:Y07241	ck: 3064	len: 141	finds: 1	! Fragment of mouse hepatic STAR
\\End of list				
Databases searched:				
Geneseq, Release 36.3, Released on 8Jul1999, Formatted on 9Jul1999				
Total finds: 245				
Total length: 23,686,106				
Total sequences: 188,963				
CPU time: 03:32.95				

990009 ck: 7203 len: 102 | Mouse gamma-2b chain New Immunoglobulin

1
1:

```
<(X)[1,200](L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W  
x[6](P)X(P(A)x(X(L)(S(R)x(X(V)XXX(C)x[80]  
FQYILPPFAEDLSRKDYSLICLVGNGIAHKKRCKQAODAKORRKLKSSCRNHPYLVEFSVGNDWIVA
```

P91884 ck: 718 len: 48 | Antigenic Epstein-Barr virus peptide Ide

1
1:

```
<(X)[1,200](L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W  
x[19](P)x(P(L)x(X(A)(S(R)x(X(P)xxx(S)x[13]  
RGWFCPSLICEEEDSGTSGTEPEPLGPASRRPPGLRSPSPVKRECLNG
```

P94157 ck: 2829 len: 246 | Acetyl-CoA reductase. Constructing new p

1
1:

```
<(X)[1,200](L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W  
x[112](L)x(S)(L)x(X(V)(T(K)x(I)x(XXX(A)x[118]  
MQRIRATVYGWGOGICTACCOLANDGFPMVAGCGENSRREKWLDEQKALGFDPIASBGNAVA  
NVTPSPGYIATDMVKAINDVDLRKIIVAIIPKRRLGPEELINISICAMLSSEEGST  
GADPSLNGDLHG
```

P90460 ck: 560 len: 288 | Alpha-factor profilinase from Southern

1
1:

```
<(X)[1,200](L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W  
x[149](L)x(T)(V)xx(V)(S(H)x(L)xxx(G)x[123]  
MRPSIFTFATFEAFSAALAPAVPTTEDTAOIPBAVTVLEGGDEOVAALPFSNSTNNLT.  
HDNAQLITAIIDGDTVLGAYGVAGCMCKSHSGVGDHSANLNLVALMAIELHN  
LCMHNDGNCHGCANSVCWAAMLSDPSCRKFSDCSKDYOTFYVNNOCIINPK
```

P81999 ck: 4995 len: 25 | Binding peptide 2 capable of binding to

1
1:

```
<(X)[1,200](L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W  
x[2](IX(P(A)xx(V)(S(R)x(X(V)xxx(A)x[7]  
CNLPASTVSRYNIYTTRQPNODIA
```

P70477 ck: 9828 len: 194 | Sequence encoding human respiratory syncytial

1
1:

```
<(X)[1,200](L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W  
x[83](I)x(S)(I)xx(I)(T(K)x(X(A)xxx(M)x[95]  
MSRRNPCKEFIEINHCHNGRCRHFSSHNYFEMPFPALLVRONFMNLRIKLSMKDSIDLTSISIGA  
ESTVSPDINDHAKNDDTT
```

P70785 ck: 9828 len: 1 | Sequence encoding human respiratory sync

1
1:

```
<(X)[1,200](L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W  
x[83](I)x(S)(I)xx(I)(T(K)x(X(A)xxx(M)x[95]  
MSRRNPCKEFIEINHCHNGRCRHFSSHNYFEMPFPALLVRONFMNLRIKLSMKDSIDLTSISIGA  
ESTVSPDINDHAKNDDTT
```

R24296 ck: 3600 len: 384 | Regulatory protein Vans involved in glyco

1
1:

```
<(X)[1,200](L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W  
x[84](V)x(S)(I)xx(L)(C)(R)x(X(L)xxx(A)x[284]  
LVYLNKNKNNDYSERKRYXMYIVALVAVLYFVILRSIMRGKLGDWILSLLENKYDLNHLD  
SLIDEAPDPKPVDOOKARYHTILDKRAYRLEOLIDEFEETFRYNLOGITILTKEHIDLY  
XYLVQMTIDEEFYPOLSAHGKQAVITHAPELDLVSGDPDKLARVFNNILKMAAYSSEDNSIIDTAGLSGDVVISIEFKNTGISIPKRLK  
R10995 ck: 1620 len: 398 | Xenopus Bone Morphogenetic Factor BMP-2A
```

1
1:

```
<(X)[1,200](L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W  
x[118](V)x(P(A)xx(G)LVPEEBKRRYSSSTRSSPOOSOVLDQFELLRNMGFLGR  
MVAGISLLILOVOILLSSHWESFEDPATIRTNIAHKRPKHGCVAVT  
IKYPAAPAAASRGCFVVELDTRLIHNEHSWESEFIDPATIRTNIAHKRPKHGCVAVT  
HLDNDITNVPKRHRVRSRSLTLDKGMPRIPLVTFFSDGGAHLKKRCKQAODAKORRKLKSSCRNHPYLVEFSVGNDWIVA
```


R24190 ck: 5872 len: 186 ! Bovine RSV strain A 51908 W2 protein. Bovid
1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{183}(L)x(S)(I)xx(I)(T)(K)xx(A)xxx(M)x(87)
MSRRNPKCYELRCHGHNKKCFHSNNYFEMAPHALVNONFMNLTKLSMDNRNDTLSEIGAAEH
DINDEONE
1:
P81092 ck: 3060 len: 95 ! Sequence of rhinovirus HRV2 viral protein F
1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{127}(L)x(P)(V)xx(I)(S)(K)xx(I)xxx(L)x{152}
GVTDYIHLMGEAFNGGFVDYKEHITHAINVGNISKRITIKMLRIISAMVIILNSSDPQTIILA
1:
P81093 ck: 5281 len: 323 ! Sequence of rhinovirus HRV2 viral protein F
1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{76}(L)x(T)(L)xx(U)(S)(R)xx(L)xxxx(A)x{1231}
SDSMKKEFEACNAARGLIEMLIGNIKSFETIPMKSLPMOAKLVKLNELKLNLEYKOVESLARAVD
GDMMTEFCOMSSVFPEPPMADLPCKGAFFSRVLCTSHSLPLPTISTPAN
RRFLFDLDIVHDNEFPDGKRLNVAAFRPDVDRIGNARKCEPFVCRAVSFKDRNSCNKYSLAOYYNMIEDRRRODVDTAII
R10974 ck: 3836 len: 246 ! Acetatecyt CoA reductase enzyme. Construct
1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{112}(L)x(S)(L)xx(V)(T)(K)xx(I)xxx(A)x{118}
MTORIAVTGGMGIGTAICORLANDGGRVAGCGPNRPERRKMLEOKKALGYDFIASSEGVAWD
1:
VNTPSPGITATDWVKAIRODIVDKIATAIPTVRKGLPREELASICAMLESEESEGST
GADFSLNGGLHMKG
R12468 ck: 7673 len: 248 ! Tauffa cylindrica bioactive protein. New pro
1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{37}(L)x(S)(A)xx(A)(S)(R)xx(L)xxx(S)x{195}
DYRSLSGSSTSYSKFIIDLRALKAPENGTAYNTLLISSAGSARYTMLTSNDTGKAIRIYAADV
1:
GVPSLATISLENISMALSJKIOLOLAQTNGRTFTPVITDDKQGORELTNYTSKVY
TRNIOLLINYKONVA
R13252 ck: 528 len: 247 ! Murine Cytotoxic Cell Protease-1. DNA vecto
1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{6}(L)x(T)(L)xx(A)(S)(R)xx(A)xxx(I)x{225}
MKILLILTLASTTAGELIIGHGHVPHSRYALLSIKDOPBEAICGFLIREDFYLTRAACE
1:
KNRNKTNOICAGDPKTRKASFRRDSGPVCRKVAAGVISTEIKODSGPARPAKFVY
SPFLSWIKTKMS
P60645 ck: 2381 len: 292 ! Mouse kidney cell Band 3-like anion transp
1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{106}(A)x(T)(V)xx(V)(T)(H)xx(A)xxx(B)x{1170}
VDSIEDITYORTKSVPSCGFVTAPADKRWGYNLGERTFEFVMANVASLLRALVVFLIMETOIT
1:
VTSINGOGEPREHLMPKHOPDVTVYKRVETRMHLEFALOCLCLLMAWNS
TAASLAPFILILLYLRLVJLIIFTEREKMCILDANAEEVFEDEGVEDENMEMPVY
R14247 ck: 3078 len: 149 ! HCV-T (6110-6557) encoded epitope. New nucl
1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{115}(L)x(S)(V)xx(A)(C)(K)xx(P)xxx(A)x{118}
VLESKAKASTVAKLLIVEACKLTPPHARSRKFGYGADVSSLSPRAINHIHSWMOLDEDDET
1:
R20004 ck: 4995 len: 25 ! VP6 inner capsid epitope, 84 TS, peptide B.
1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{2}(L)x(P)(A)xx(V)(S)(R)xx(V)xxx(A)x{7}
CNIAPASIYSKIIVYTROPQDI
1:

R21409 ck: 617 len: 482 ! NADH dehydrogenase subunit 2. New DNA se
1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
x{184}(L)x(S)(L)xx(F)(S)(K)xx(I)xxx(X)x{1282}
MLLSIIISOALASSMSMYVALLSRIISITIELLYNNVEIIGLGATGIYNFLOYT
1:
SVTGTISESAISKYNLINYOISLICYGLFKIGFYHPHONRIDVIDPOVP
IITWTLTITKLSILPIEMFIVHSESSEMWTIIMLTSLVAVGLSIGOSRIKRLIYSNWSHVGLMLSITEKSEAEF
R23004 ck: 7059 len: 297 ! Protein transcribed from the mba sequenc
1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
x{123}(A)x(S)(V)xx(I)(T)(R)xx(L)xxx(I)x{258}
MDPLLKKLFKITELMETGESFSIASIIVTYRTPSRISVLSBELKCOLFIKRGKIIIPREF
1:
GPCDLINLAGPYVPHGANNFNLDIHYEFFEQTLINPAPSFDNDVLFSSYRL
QQCALMLIPVNCALGLSIDBALHKVVALCTSIFYPTKKRETPIYRNAIKLIOELKOSTF
R22695 ck: 4919 len: 286 ! blaC31 mutation (5). Novel DNA modifiest
1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
x{193}(A)x(S)(V)xx(A)(S)(R)xx(L)xxx(M)x{77}
NETOFERALLYPEFAACLPVFAHPTELIXKVADADOLGARVGYLELMSGILIESFERPER
1:
TTMFAPAATUTLKRLASSVTLASSTLASRCOODLROMMEADRAGVLLRSDL
GAGERGSRGITALGDPGPRFIVYITIGSQATNDBERNROIAGEIASLIKHM
R25601 ck: 5700 len: 85 ! NANBH virus antigen. Antigenic polypept
1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
x{32}(L)x(S)(V)xx(A)(C)(K)xx(P)xxx(A)x{37}
KRVTEKRLQVLDHDIRIDVLEKEKAKASYKAKLLSVEACKLTPPHARSRKFGYAKDVNRN
1:
R25869 ck: 3783 len: 239 ! HCV polyprotein 16. Hepatitis C virus an
1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
x{84}(L)x(S)(I)xx(A)(C)(K)xx(P)xxx(B)x{139}
VCCMSYTWGALTIPCABESBKLDELINSLLRHSHSVITSBSASLRKKTVEPRLQVL
1:
LYDAVSTLPGAWMPSTFYQSFGQRVEFLVNTWKSKCPMGESTIDRFCDSTVTE
NDIRE
R25891 ck: 4644 len: 253 ! HB16. Hepatitis C virus antigen expresse
1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
x{96}(L)x(S)(I)xx(A)(C)(K)xx(P)xxx(B)x{141}
MITPSLBHCMOVOCSSWYSTWGALTTPCAAESBLPLNPUSNLSLRHSHSVITSBASLR
1:
PDIGVFCCKALTYNSTLPGAWMBSTFYQSFGQRVEFLVNTWKSKCPMGEST
YDRCDFSTVTEINDIRITGL
R25303 ck: 9668 len: 194 ! HRSV 22K protein. Vaccines for human res
1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
x{83}(L)x(S)(I)xx(I)(T)(K)xx(A)xxx(I)x{95}
MSRTNCAFETIRGHOLNKGRKCHSHNFEMPHALVNONFMLNRINKLSMDKSIDTLSELISAQA
1:
ESTVSDTNDAKNNDDT
R29874 ck: 9846 len: 171 ! HCV NS4-NS5 peptide N29-1, N29-2, N29-3
1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
x{18}(L)x(S)(V)xx(A)(C)(K)xx(P)xxx(A)x{137}
YRDVLKEMKASTVAKLLSVEACKLTPPHARSRKFGYGADVSSLSPRAINHIHSWXDXD
1:
R29879 ck: 8945 len: 285 ! HCV NS4-NS5 peptide 2918. New hepatitis
1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
x{18}(L)x(S)(V)xx(A)(C)(K)xx(P)xxx(A)x{251}
YRDVLKEMKASTVAKLLSVEACKLTPPHARSRKFGYGADVSSLSPRAINHIHSWXDXD
1:

OCDDLPDEARQAIRSLTERLYIGGPLINSKQONCGYRCRASGVLTSCGNLTLCY
LKASACRAKIQDCTMLVCGDDLVVICESAGTQEDAPANLRFTEAMTRNSA

R29895 ck: 8862 len: 171 1 HCV NS4-NS5 peptide N29-1. New hepatitis C

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{18})(L)x(S)(V)xx(A)(C)(K)xx(P)xxx(A)x{137}
1 YRDVLKEMAKASTYKAKLISVEEACKLTPHSAKRSKFGYGAADVRLSSKAVNHRISWKDLLE

R29896 ck: 8862 len: 171 1 HCV NS4-NS5 peptide N29-2. New hepatitis C

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{18})(L)x(S)(V)xx(A)(C)(K)xx(P)xxx(A)x{137}
1 YRDVLKEMAKASTYKAKLISVEEACKLTPHSAKRSKFGYGAADVRLSSKAVNHRISWKDLLE

R29897 ck: 2886 len: 167 1 HCV NS4-NS5 peptide N29-3. New hepatitis C

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{18})(L)x(S)(V)xx(A)(C)(K)xx(P)xxx(A)x{137}
1 YRDVLKEMAKASTYKAKLISVEEACKLTPHSAKRSKFGYGAADVRLSSKAVNHRISWKDLLE

R29909 ck: 429 len: 277 1 Prod. of the luffin-f gene. New gene coding

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{57})(L)x(S)(A)xx(A)(S)(R)xx(L)xxx(S)x{204}
1 MKRFVLLIATVVAADYRELSGSSSTYSKFGIDLRKALPENGTVYNTLLSSAGAST

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{57})(L)x(S)(A)xx(A)(S)(R)xx(L)xxx(S)x{204}
1 MKRFVLLIATVVAADYRELSGSSSTYSKFGIDLRKALPENGTVYNTLLSSAGAST

R29910 ck: 9157 len: 278 1 Prod. of the luffin-g gene. New gene coding

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{57})(L)x(S)(A)xx(A)(S)(R)xx(L)xxx(S)x{204}
1 MKRFVLLIATVVAADYRELSGSSSTYSKFGIDLRKALPENGTVYNTLLSSAGAST

R31608 ck: 7012 len: 205 1 Homologous to chicken nov gene exon 3-4-end

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{110})(L)x(P)(A)xx(P)(S)(R)xx(L)xxx(V)x{79}
1 QIFRIPLDADVRPOCLTSASPTLFPSSPAKDGACLEGGIVYNSGEFSOCSCKYQCTCLDCH

R30618 ck: 7998 len: 106 1 Polypeptide coded by Korean HCV cDNA fragment

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{54})(P)x(P)(I)xx(I)(T)(K)xx(L)xxx(G)x{36}
1 LRLTSPYKVFARLILMWLQYFITRAEHLQWIPPLNRGSRDAILLACAVAPPEIPITKX

R32192 ck: 2829 len: 246 1 Sequence encoded by the acetosetyl-CoA red

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{112})(L)x(S)(L)xx(V)(T)(K)xx(I)xxx(A)x{118}
1 MTORLAVYVGGGGIGTAICORLADGFRVAVAGGPNSPRERKLEDOOKALGDFDTASGNVADMC

R37294 ck: 7673 len: 248 1 Plant type I RIP luffin. Analogues of type

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{37})(L)x(S)(A)xx(A)(S)(R)xx(L)xxx(S)x{195}
1 DVAFSLSGSSSTYSKFGIDLRKALPNSGTVYNTLLSSAGASRYTLMTLSNYDKAIVAVDP

TKNIQILLNKNVNA

R39259 ck: 8652 len: 391 1 Human somatostatin receptor-1. Somatosta

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
x{188})(I)x(P)(I)xx(F)(S)(R)xx(A)xxx(G)x{187}
1 MPFNGTASPSFSSPSGSGCEGAGSGADGMEFGRNAONGTISEGCGSAILISF

R39260 ck: 8110 len: 391 1 Murine somatostatin receptor-1. Somatost

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
x{188})(I)x(P)(I)xx(F)(S)(R)xx(A)xxx(G)x{187}
1 MPFNGTASPSFSSPSGSGCEGAGSGADGMEFGRNAONGTISEGCGSAILISF

R42317 ck: 3489 len: 345 1 EBV vCA-p40 from ORF BHRF1. Peptide(s) a

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
x{51})(A)x(P)(L)xx(F)(S)(H)xx(I)xxx(P)x{278}
1 MLSGNAGEATACGSAAGADLISVPRNTFTLQTLNDLRKPRQTPPLPYAPLPFSSHQAI

R52667 ck: 5723 len: 303 1 Equine herpesvirus US2 gene. New recomb1

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
x{62})(L)x(T)(L)xx(A)(C)(R)xx(G)x{125}
1 MGVLITVTVVNDNRHALLNSIDVDGLMELTSCQCVLSEFLGPIVYVRSADLYRFSSTL

R49296 ck: 4901 len: 25 1 Invariant chain (I1) position 97-121. NO

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
x{P)x(P)(P)xx(V)(S)(K)xx(M)xxx(L)x{7}
1 LPKPPKPYSKKRNATPLLMQALPM

R49297 ck: 3126 len: 24 1 Invariant chain (I1) position 97-120. NO

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
x{P)x(P)(P)xx(V)(S)(K)xx(M)xxx(L)x{7}
1 LPKPPKPYSKKRNATPLLMQALPM

R49299 ck: 1278 len: 23 1 Invariant chain (I1) position 97-119. NO

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
x{P)x(P)(P)xx(V)(S)(K)xx(M)xxx(L)x{7}
1 LPKPPKPYSKKRNATPLLMQALPM

R49438 ck: 4300 len: 49 1 Murine HLA-DRA alpha chain/IT 24 residue

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
x{26})(P)x(P)(P)xx(V)(S)(K)xx(M)xxx(L)x{7}
1 MAISGVPLGFFIIVAVMSAOESWALPDKPKPYSKKRNATPLLMQALPM

R49589 ck: 4300 len: 49 1 Sequence of HLA-DR alpha chain leader pe

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
x{26})(P)x(P)(P)xx(V)(S)(K)xx(M)xxx(L)x{7}
1 MAISGVPLGFFIIVAVMSAOESWALPDKPKPYSKKRNATPLLMQALPM

R49589 ck: 4300 len: 49 1 Sequence of HLA-DR alpha chain leader pe

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
x{26})(P)x(P)(P)xx(V)(S)(K)xx(M)xxx(L)x{7}
1 MAISGVPLGFFIIVAVMSAOESWALPDKPKPYSKKRNATPLLMQALPM

R49589 ck: 4300 len: 49 1 Sequence of HLA-DR alpha chain leader pe

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
x{26})(P)x(P)(P)xx(V)(S)(K)xx(M)xxx(L)x{7}
1 MAISGVPLGFFIIVAVMSAOESWALPDKPKPYSKKRNATPLLMQALPM

[illegible]

W04831 ck: 1679 len: 207 1 Vascular endothelial growth factor-B186. Vd
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
 x{158}(A)x(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
 1: MGLRLRLALLQLAFAQVSPDPAHQKRVSNIDVITATQCPREYVPLVELMGIVAK
 AAPSTTSALTPEPAAAAADAAASVAKGA
 W11481 ck: 5195 len: 205 1 D. Immitis mature venom allergen antigen 5-
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
 x{157}(A)x(T)(L)xx(I)(C)(H)xx(P)xxx(M)x{32}
 1: YCEGGGKLPTEKRNIVQINKYRSRLRGRLKKNKGDIAPKRNMDLRMRDCKLEKSAQNMAMNC
 IYELGNPCRNKNDCKTRKCSAKSGICKK
 W11479 ck: 9389 len: 221 1 D. Immitis venom allergen antigen 5-like pr
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
 x{173}(A)x(T)(L)xx(I)(C)(H)xx(P)xxx(M)x{32}
 1: MLLVTFPAILIYVVAAYCEGGKLPTEKRNIVQINKYRSRLRGRLKKNKGDIAPKRNMDLRMRDCKLEKSAQNMAMNC
 VFICHPFGGMVNDLIYELGNPCRNKNDCKTRKCSAKSGICKK
 W13272 ck: 2018 len: 244 1 Rhodococcus erythropolis SK92-B1 regulatory
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
 x{179}(A)x(P)(W)xx(F)(S)(R)xx(L)xxx(V)x{49}
 1: MAGADHAAGGNRRARILVDDKHFRTVYTWQLESENFPVVAADGADALAKQVTEGADLMVLL
 MAASPMQVFSRRRLLEVRSSPMDQDVAITEVHRIKRIEDPFPKPIILQITVR
 GAGYRFDEGRA
 W06943 ck: 9316 len: 140 1 Cagi locus product 14. Helicobacter pylori
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
 x{40}(L)x(S)(V)xx(V)(S)(R)xx(L)xxx(I)x{84}
 1: KRIFIKNTYLLGVSLACIOSVNADONTDIKDISPEDMALNSGLVSRDQLKIEIPKETLEOKVA
 W06935 ck: 8017 len: 136 1 Cagi locus product 6. Helicobacter pylori
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
 x{102}(L)x(S)(L)xx(I)(S)(R)xx(L)xxx(A)x{18}
 1: NIYFELLYKFTALNTYFRNONGINQISKLONFLQKRYISFNKHLDKSLYKLFNISTIVIGFLI
 W20571 ck: 6963 len: 114 1 H. pylori secreted or periplasmic protein 8
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
 x{80}(L)x(S)(L)xx(I)(S)(K)xx(L)xxx(A)x{18}
 1: MKRIRISKLONFLQKRYISFNKHLDKSLYKLFNISTIVIGFLIALFSYGAVILVPIELFA
 W20886 ck: 2872 len: 154 1 H. pylori secreted or periplasmic protein,
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
 x{54}(L)x(S)(V)xx(V)(S)(R)xx(L)xxx(I)x{84}
 1: NNAKKNKGAVKMTNFYKIKLFAFWCLIIIGMFANPLNADONTDIKDISPEDMALNSGLVSRDQL
 W20776 ck: 6185 len: 276 1 H. pylori flagella-associated protein, 0796
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
 x{45}(L)x(S)(I)xx(I)(S)(K)xx(V)xxx(G)x{215}
 1: RLKMAKLPKQKADDELMSSEKIAIILIOVGEDTGEILRLHLDISTEISKOIVQLNGTDKQI
 NGEISPOVYKRVSTLENKLESTLESYKRAAEIENRIGOKSAKTTIARI
 ESDVNDKLAGAIKEMFTEIDIAKLDNFAMKDFKSGGIKRTGL
 W20259 ck: 5141 len: 108 1 H. pylori secreted or periplasmic protein 2
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
 x{173}(A)x(T)(L)xx(I)(C)(H)xx(P)xxx(M)x{32}

W20324 ck: 6372 len: 102 1 H. pylori cell envelope transporter prot
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
 x{2}(A)x(S)(L)xx(V)(S)(R)xx(L)xxx(L)x{84}
 1: MMAHSLIVSKTSLSNLIIFVOPDGKLSMTDAIDPMTNGLRMVYNEIAEKFKLLND
 W20696 ck: 7893 len: 121 1 H. pylori secreted or periplasmic protei
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
 x{87}(L)x(S)(L)xx(I)(S)(R)xx(L)xxx(A)x{18}
 1: TSSFRTKMRPISKLOKONFLQKHSFNKHLDKSLYKLFNISTIVIGFLIALFSYGAVI
 W20139 ck: 1367 len: 53 1 H. pylori cytoplasmic protein, 14257751.
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
 x{33}(L)x(T)(L)xx(I)(S)(H)xx(V)xxx(P)x{4}
 1: MLDPLVLEFATGDLAMKRLFVSLYEIISTWVLKTLIGLSHGRVSPMKSF
 W09433 ck: 8653 len: 328 1 Human placenta purinergic P-2u receptor,
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
 x{175}(L)x(P)(P)xx(A)(T)(H)xx(P)xxx(A)x{133}
 1: MEMWNGTDQALGPPTTCVARENKOLLEPVSALAPALPLNCIVITQITSRRALITRAV
 YDLSPPALATHMYEGNALTIVIFLFPALLACVCLCRICRQDQDEPVAQDER
 RGKARAAVVAAYFGLIFLPHITATVAVSTPGVCTLEAFVAAYKGRFPASANSVLDPILEYFTQKKRRRPHELLQK
 W21704 ck: 7673 len: 248 1 Luffin-A. Inactive precursor of maize r1
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
 x{37}(L)x(S)(A)xx(A)(S)(R)xx(L)xxx(S)x{195}
 1: DVRSLSGSSSTYSKISGILKRALPSKNGVYVLLTLLSSAGASRYTLMTLSNDGKAITVA
 QVPSLATISLNSLSKQIOLAQTNGTFTPVITIDRGKQVEITNVTSKV
 W21709 ck: 7681 len: 250 1 Luffin-B. Inactive precursor of maize r1
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
 x{38}(L)x(S)(A)xx(A)(S)(R)xx(L)xxx(S)x{196}
 1: ANVSFSLSGSDSKISYKFTALRKALPSKESNIPLLPSASGASRYTLMTLSNDGKAITVM
 QIKDVNSKLLINKONIA
 W10031 ck: 8373 len: 74 1 Protein encoded by clone G-4-5 from act1
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
 x{8}(V)x(P)(P)xx(A)(S)(R)xx(P)xxx(P)x{50}
 1: WILRLRPVHHPQGASRREPRLPOGRRRRQVLLASIVIASLPIEIVCMGCVGCHLPT
 W13647 ck: 5362 len: 30 1 Invariant chain region which contains th
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
 x{P}(P)(P)xx(V)(S)(R)xx(L)xxx(L)x{13}
 1: LRFPRFVSKMKAIPLLMQALPMGALPQGS
 W25141 ck: 7673 len: 248 1 Luffin-A (a ribosome inhibitory protein)
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
 x{37}(L)x(S)(A)xx(A)(S)(R)xx(L)xxx(S)x{195}
 1: DVRSLSGSSSTYSKISGILKRALPSKNGVYVLLTLLSSAGASRYTLMTLSNDGKAITVA
 QVPSLATISLNSLSKQIOLAQTNGTFTPVITIDRGKQVEITNVTSKV
 TKNIOILLNKKONVA

1	W50006	ck: 5567	len: 365	1	Human papillomavirus-18 E2 protein. Treated
					<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C x{61}(V)x(T)(A)xx(A)(S)(R)xx(L)xxx(L)x{288} x{187}(A)x(T)(L)xx(A)(S)(R)xx(L)xxx(L)x{104} MKRVLLTSLSLFEMLAERNGFYLGNFAEGSYIGOGSGIGEKASAEANALNAINNAKNSLFPK YKDMTGRLLDADTLKRAGRHIIKRSQGLVIGMEIGASTWFAFNSNLTAFENOVKSGRTI FOLQCKFGVRRSSDEYDIDRYGDENYIGSSVELGVKVPKVVYSDDYGDLDKRVVSVYLYNTYFKKHK
1	W55590	ck: 428	len: 307	1	H. pylori ORF 09cpi0224_429510_c2.46.na cel
					<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C x{187}(A)x(T)(L)xx(A)(S)(R)xx(L)xxx(L)x{104} MKRVLLTSLSLFEMLAERNGFYLGNFAEGSYIGOGSGIGEKASAEANALNAINNAKNSLFPK YKDMTGRLLDADTLKRAGRHIIKRSQGLVIGMEIGASTWFAFNSNLTAFENOVKSGRTI FOLQCKFGVRRSSDEYDIDRYGDENYIGSSVELGVKVPKVVYSDDYGDLDKRVVSVYLYNTYFKKHK
1	W55815	ck: 5054	len: 271	1	Streptomyces roseofulvus frenolicin gene cl
					<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C x{125}(L)x(S)(V)xx(V)(T)(R)xx(L)xxx(G)x{130} MTTAAPHETPFGAGTTRGPALVTGATRGIGLVAELVARGVPVVCADPAEAARTVELAAGGA TKALGLEIARSGITVNAVCPGYVETPMAGVRRHYADLMVDYEDVLAAREAKIPL GRTYRDEVALVLYLTDAAAATVTAOLNVCGLIGNY
1	W56137	ck: 7408	len: 186	1	Open reading frame 2 peptide of bacteriophage
					<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C x{74}(L)x(P)(M)xx(P)(S)(R)xx(A)xxx(A)x{196} MNPEDLNLTSNVSAGVAGSLFELPGLASGVLDIGAKONAKOJARQMAFOERMSTAYORA KULFSRKGR
1	W51011	ck: 7216	len: 277	1	Human liver carbonyl reductase. DNA encoded
					<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C x{189}(L)x(P)(M)xx(P)(S)(R)xx(A)xxx(A)x{172} MSSGILAVLTGNGNGIGALIVRDLCLFSGDVLADYADVTGQAAVQOQAAGLSPTFHQDLIDI DTRKGVHOKEGMPSAGVYGTGIVTSIRHARLGRKDKILNACCPCGWVRT DMAGKATKSPSEGAETPVYLAALLPDAEPGHGOFVSEKRYEQW
1	W28052	ck: 2273	len: 269	1	Amino acid sequence of phospho-beta-glucosyl
					<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C x{43}(L)x(P)(V)xx(L)(S)(H)xx(M)xxx(L)x{210} MGLKCLRTSIATRTFPGDEDEVPNEDEVLADYDRFELLAQGLEPVTLSHEMPLHLAKHYGF AOIANRLFEFFPDVQVGRGYPSYAKRMIAKRGYDVGQEDDSLTDOGTXYDITGS YMTSAVHVHDVDTXENNIVNGGFCIGESAYENE
1	W37816	ck: 3813	len: 317	1	Human secreted apoptosis-related protein hs
					<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C x{170}(L)x(P)(T)(A)xx(V)(T)(R)xx(A)xxx(M)x{131} MRAAAGGVTRALALDLGLALHMAKRCLEEDYTGMADEPLHGRSISKPPQCLDIPADLPICHTY TTCACQCEHNSADLMEQNCSSDFVVKMRKEIKIENGKRLIGAKKRLKLPFG PLKRRDTRKLVLMHNGAGCPCLDPLDLAGSFLVGRKVDGOLLAAVYMDKRNKEMFAVKEFMSYPCSLYFFFGAAEPH
1	W59024	ck: 9911	len: 71	1	Enterocin-900. Bacterial growth inhibiting
					<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C x{27}(L)x(T)(P)xx(L)(S)(R)xx(A)xxx(L)x{28} MOMVKEIESTKEMQOIIIGGENDHMRPELNTPTNLSKGGKCAALAGSLFIPRGPPLAMAAALA W57330
					ck: 4443 len: 339 1 Glycero1-3-phosphate dehydrtogeenase gpsa. Fe
					<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C x{68}(L)x(T)(A)xx(A)(S)(R)xx(L)xxx(P)x{255}
1	W62641	ck: 7060	len: 137	1	Fllea serine protease inhibitor variable
					<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W) x{88}(A)x(T)(A)xx(A)(T)(R)xx(L)xxx(A)x{33} BEKLQNDLQNLQRMYSVEVILDLPRKFIRESINLNDLKLKQMSDMFVPGKADFKLLEGS
1	W55052	ck: 4306	len: 118	1	Sunflower antifungal protein M559 fragme
					<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W) x{41}(L)x(T)(A)xx(V)(S)(H)xx(A)xxx(C)x{61} DPSFPIETGEVTPGSSSPFTVLQVYINLNRENTTTPKPLIITAEHVSHIQAAVYCGKON
1	W48722	ck: 2863	len: 355	1	Human V28 seven transmembrane receptor.
					<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W) x{80}(V)x(T)(A)xx(W)(T)(H)xx(L)xxx(G)x{259} MDQFPESEVTENFEVDLAEACYIGDIVEYFGLSLFYSVIRAIQVGNLVLVFAITNSKRPK DYEPVLOEIMPVLRNVEINLGLFLPLDLSMSYCFRIOTLFSCHNRKAKAIKI LLVIVFLEFWETPVNVIFFETLKYDFPSCDMRDLRLALSVETVAFSHCINPLVAFAGEKFRRLYLHLGKCLAVLCGR
1	W54728	ck: 3126	len: 24	1	Peptide from II p80-103. Increasing upta
					<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W) x{P}(P)(P)xx(P)(S)(R)xx(W)xxx(L)x{77} LPPKPKVPSKRMATPILMQALPM
1	W60258	ck: 4443	len: 339	1	Klebsiella pneumoniae glycero1-3-phospha
					<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W) x{68}(L)x(T)(A)xx(A)(S)(R)xx(L)xxx(P)x{255} MNQNSAMETVIGAGSGTALATILANGHEVYLGMDPHEIATLERDRQNAAFLPDVPFDTL RYVSNPFGVQAGVANKVYALGAGMSGIEYFAGANARALITRLAENSRLGAL GADPATMGAGAGDLVLTCTDQNSNRFRGMILGQMDVQAGCKIQGVVEGRNTRKEVRELHNRGEMPTTEETIYQVLYCGK
1	W60267	ck: 843	len: 387	1	Klebsiella pneumoniae DHAT protein. New
					<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W) x{143}(A)x(T)(A)xx(V)(T)(R)xx(V)xxx(T)x{228} MSYRMFVYLVNPNVFEQPNISVAGRCQLGKAKALVTDGLAIRDGAVDKTLHYLRAG INDPLMIGRPAALITATGMDALTHAVVEYIEKDNAPVYDAAMAQALILIRNRNO AVALGSMIQARENMAVASLLAGMAFNANMIGYHAMAHQGLGIDMPGHVANAVALPDAVRYNLIANPEKFAIDIELGEMINITGL
1	W60764	ck: 7422	len: 260	1	Rainbow trout interleukin 1 beta. Rainbo
					<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W) x{84}(L)x(T)(A)xx(A)(S)(R)xx(L)xxx(P)x{160} MDFSNSTILKNSSESANMSKRPQSLDIEVSHPTMHNHNLIIANERLKGSGVYMGTEF RPTLHLEEVADKQDLISISQSDNVFLLTFRRTGVDISTLESASFNMFTSTDMQ QDYTKPVMQCKAKAPNMLTFTIQRHN
1	W57889	ck: 5290	len: 271	1	Corn raffinose synthetase. New nucleic a
					<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W) x{1170}(V)x(P)(V)xx(L)(S)(H)xx(A)xxx(L)x{85} OSTHPCGAFAAASASISGPIYVDSVGHDBALLRRLALPBGVTLRREGNALPFRDCLFADP SHERAIKFAPIGLANMLNAGANQARAKKDCGVTAEFVYGADELVAIYASAIPIR LCRVNGDEAFETTKDGYVTVYDVYRSGSSSKLCCVOYV
1	W44905	ck: 1135	len: 39	1	"polyprolline beta-turn helix" polypeptid
					<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W) x{68}(L)x(T)(A)xx(A)(S)(R)xx(L)xxx(P)x{255}

				x(23)(P)x(T)(P)x(x(W)(S)(H)xx(L)xxx(S)x(211)) MNSMSTAFAVANSVLVVAAPHNGYFVAVPGIMSHVPLPNSQPOVHVLPGNPSPLSVNVNGOPVOKAL VNPGMALSGVLLVLEFGLIEFGIACASSHFEGCOLVCCOSSNVSIYPMIYKANPVITP EPVTSPPSYSEIQANK
1:				
W65460	ck: 1814	len: 375	i Human growth differentiation factor-8. Transg	
1:				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C x(65)(L)x(T)(A)xx(I)(S)(K)xx(I)xxx(L)x(294) MOKLOLCYITILEFMLIYAGVVDLNESEKENEKGLCNACLMRONKSSRIEAIKIQIILSKRL KPMKDGRTYTGIRSLKIDMNGPGTIGWOSIDVKTVDLONMLKOPESNLGIEIKALDEN GHDLAVTFPPGEGDLNPFLEVKYVTDTPKRSRDFGLDCDEHSTESRCRCRPLTVDFEAFGMDWIIAPKRYKANYSCEGCEFFVL
W64373	ck: 3575	len: 120	i Mycobacterium tuberculosis antigen RDIF5. N	
1:				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C x(95)(A)x(S)(A)xx(A)(C)(R)xx(P)xxx(P)x(9) TPERKSFVDDLDIDLSMWEIANGQEDKGVKIPEDDLAGLRTYGVVAYIQKLEENPEAQAAL
W63580	ck: 5404	len: 174	i Streptococcus pneumoniae protein of unknown	
1:				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C x(72)(L)x(T)(V)xx(L)(S)(K)xx(V)xxx(I)x(86) MGFNOLPSMLILHAKGLTESFIKASAGTADRTITYSCHVRNHGAVDIKLYOELGELKRYQPYDI
W70962	ck: 4556	len: 212	i A STAT function regulatory protein designat	
1:				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C x(171)(V)x(P)(L)xx(L)(C)(R)xx(I)xxx(V)x(25) MYARNOVAADNAISPALEPRRSPSSSSSSPAAPVPRCPGVPAPABGDTFRIFRSHSDYF LCORRYAANGRENLIARIPLPVLRDILSSFPQI
W77666	ck: 5825	len: 100	i Staphylococcus aureus protein of unknown fu	
1:				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C x(176)(A)x(T)(P)xx(A)(S)(K)xx(I)xxx(I)x(8) VMPXVFWMPFXIFIGTLPLAGVATGAIHFTANEVPIPGXMLXNNGLAITLAVLQNDRAFY
W38716	ck: 3773	len: 161	i S. pneumoniae aspartyl tRNA synthetase. Nov	
1:				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C x(53)(P)x(T)(A)xx(I)(S)(R)xx(I)xxx(T)x(92) MKYDDAMALYGSKDPDTRFMDLLODLTEVYKGVDFYSEALAVKAIYVNGLEPLADNYSRKDID
W79302	ck: 8790	len: 67	i A Staphylococcus aureus protein. New nucle	
1:				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C x(50)(L)x(S)(A)xx(I)(T)(K)xx(I)xxx(M)x SHNLNRGAFLAVIGDILGSLGVAITRALIAPFMTIADPIASILVSVILKRAMTITSSINI
W69889	ck: 1496	len: 376	i Rat growth differentiation factor-8. Transg	
1:				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C x(66)(L)x(T)(A)xx(I)(S)(K)xx(I)xxx(L)x(294) MIOKROMVYIYLLFVLIAAGVVDLNESESEAVNEEGGLCNACLMRONKSSRIEAIKIQIILSKRL IKPMKDGRTYTGIRSLKIDMNGPGTIGWOSIDVKTVDLONMLKOPESNLGIEIKALDEN NGHDLAVTFPPGEGDLNPFLEVKYVTDTPKRSRDFGLDCDEHSTESRCRCRPLTVDFEAFGMDWIIAPKRYKANYSCEGCEFFVL
W69891	ck: 1805	len: 375	i Pig growth differentiation factor-8. Transg	
1:				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C x(65)(L)x(T)(A)xx(I)(S)(K)xx(I)xxx(L)x(294) MOKLOIYVYITILEFMLIYAGVVDLNESEKENEKGLCNACLMRONKSSRIEAIKIQIILSKRL KPMKDGRTYTGIRSLKIDMNGPGTIGWOSIDVKTVDLONMLKOPESNLGIEIKALDEN

1	W82504	ck: 447	len: 225	Human EPRG1 protein #1. New EPO primary res
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				PQHGRKRVSMIDYTRATCOPREVYVPLVELMGTVAQVAVQVSCVTVORCGGCCPDGEGCEVP
1	W82505	ck: 9756	len: 157	Human EPRG1 protein #2. New EPO primary res
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				KVSMWIDYTRATCOPREVYVPLVELMGTVAQVAVQVSCVTVORCGGCCPDGEGCEVP
1	W81740	ck: 3575	len: 120	M. tuberculosis immunogenic polypeptide RD1
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				TPKFSVDDLDIDSLMVEIAVQTEKRYGKIPEDDLGLTVGDVAVATQKLESENEAQAQ
1	W72902	ck: 5027	len: 226	Mycobacterium tuberculosis antigen CP19A.
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				MIPROPQSHGRWRAAGARLITSLVAAFAATLLPLALPFRASAGCCDAEYVFAARGGEPGIGF
1	W80495	ck: 1679	len: 207	Human vascular endothelial growth factor (V
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				MSPLRLRLALLDLAPAOAVSQDPDFGHRKRVSMIDYTRATCOPREVYVPLVELMGTVAQ
1	W85306	ck: 3126	len: 24	Helper T-cell class II peptide derived from
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				LPKPKPVSKNRMATPLLMQALPM
1	W86202	ck: 9273	len: 206	Human VEGF-related factor (VRF)-2 sequence
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				MSPLRLRLALLDLAPAOAVSQDPDFGHRKRVSMIDYTRATCOPREVYVPLVELMGTVAQ
1	W86541	ck: 8848	len: 420	Amino acid sequence of the SIAx DP2-64 (Oct
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				MSPLRLRLALLDLAPAOAVSQDPDFGHRKRVSMIDYTRATCOPREVYVPLVELMGTVAQ
1	W86235	ck: 8227	len: 185	Human VRF (VEGF-related factor)-2 full leng
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				PVSGPDAPGHOKRVSMIDYTRATCOPREVYVPLVELMGTVAQVAVQVSCVTVORCGGCCPDG
1	W86214	ck: 5797	len: 178	Human VRF-2 truncated fragment 1. New trunc
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				PQHGRKRVSMIDYTRATCOPREVYVPLVELMGTVAQVAVQVSCVTVORCGGCCPDGEGCEVP
1	W86215	ck: 4363	len: 173	Human VRF-2 truncated fragment 2. New tr
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				KVSMWIDYTRATCOPREVYVPLVELMGTVAQVAVQVSCVTVORCGGCCPDGEGCEVP
1	W86216	ck: 2203	len: 168	Human VRF-2 truncated fragment 3. New tr
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				IDYTRATCOPREVYVPLVELMGTVAQVAVQVSCVTVORCGGCCPDGEGCEVP
1	W86217	ck: 3123	len: 163	Human VRF-2 truncated fragment 4. New tr
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				RATCOPREVYVPLVELMGTVAQVAVQVSCVTVORCGGCCPDGEGCEVP
1	W85119	ck: 5663	len: 446	A delta-5 desaturase enzyme. New nucleic
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W73507
1				ck: 3945
1				len: 317
1				Human ATG-1709 protein. New human secret
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				MRAAAGAGVETALALLGALMAAPARCEHYGWOQEPHGHGSKPPOCCDIDRADPJC
1				TKICACCEHESADGIMOMOSDFVYKMKIKETIENDGRKLGAKOKKKLLRG
1				FLNRKDKLVYHMKNGAGCCPDGLASFLVMGRKVDGQILLMAYVMDKKNKEMFAVEMFSPCSLYPFYGAAPH
1	W90011	ck: 9038	len: 343	Expressed antigen for cluster 35a. New H
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				MATKLPKOKAOLDELMSERIALILOVABDTGETLRLDIDSLTEISKQIVGLNGTDKROI
1				FIPOVYKRVSTVLENKESLISYKIEVGLRAVAEIRNIGOKRSKRTTLARIESV
1				DNKLGAIKEMMTFEDYKLDNFARILKVAADKDLISLAKTSTRLDQFLNNMSRAAEQVFEEMOVLGAVKIKVDVAQR
1	W89910	ck: 8829	len: 343	Antigen 1 from cluster 35a. New Helicoba
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				MATKLPKOKAOLDELMSERIALILOVABDTGETLRLDIDSLTEISKQIVGLNGTDKROI
1				FIPOVYKRVSTVLENKESLISYKIEVGLRAVAEIRNIGOKRSKRTTLARIESV
1				DNKLGAIKEMMTFEDYKLDNFARILKVAADKDLISLAKTSTRLDQFLNNMSRAAEQVFEEMOVLGAVKIKVDVAQR
1	W85590	ck: 2306	len: 354	Polypeptide with 98% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1				x{1121}(A)X(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
1				W85591
1				ck: 8607
1				len: 351
1				Polypeptide with 95% homology with human
1				<(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
1		</		

[illegible]

W97840 len: 207 1 Winged bean chymotrypsin inhibitor WCI-3. A
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
 x(59)(V)X(S)(P)XX(V)(S)(K)XX(P)XX(S)x(122)
 MKSTTFALFLSLAISLSSTADDDVDAESNINVENGGYLLPPIHMHGGGIEFTATGNEPCH
 1:
 RNGNRRLVTEENLELVLLKAKSETASSH

Y03794 len: 144 1 S. aureus polypeptide. New essential polype
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
 x(117)(L)X(T)(L)XX(I)(T)(K)XX(V)XXX(G)x(111)
 MDRTKSLINVFVGNKRALDLEQITKEVDKRYGLWITERAVLELLYKNGQPPQIRIDRVLLIASSS
 1:
 Y03782 len: 135 1 S. aureus polypeptide. New essential polype
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
 x(92)(I)X(S)(I)X(L)(T)(R)XX(L)XXX(S)x(27)
 IHERAFIKRGLDLPADNGFQCGIWEGLTVGLGLHETIMHAKKISLGYLYKKQEGHGIMTQF
 1:
 Y02934 len: 53 1 Fragment of human secreted protein encoded
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
 x(3)(A)X(S)(P)XX(L)(S)(R)XX(M)XXX(C)x(34)
 TLVAGSPCSLSRWIMAGFCHGELVQSDMEQEMERGQVILSHSLPWCYVSPR
 1:
 Y11970 len: 113 1 Human 5' EST secreted protein SEQ ID No: 57
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
 x(9)(V)X(P)(L)XX(L)(S)(R)XX(L)XXX(C)x(88)
 MANLFIKRVNPLLYISRHIVKRALSTLFGSINGAAPVAVEGAAVRSLSLSPGLPHILPRLI
 1:
 W96262 len: 423 1 Brn-3a polypeptide. New polypeptide compris
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
 x(136)(M)X(S)(L)XX(L)(S)(H)XX(A)XXX(M)x(211)
 MMSRNSQPFPAHAPLIPERKXPSLSHSESEIRACLETPLQSNLFASUDELTLARABALAVIDI
 1:
 GGGGPGGGLGSSAHPHMSHSLGHSLSHRAKAAANMPCSPHGLVAAAHHGAA
 AAAAAAGGVAAAASAAVGAAGLASICDSDTDFRELEFAERFKORIKIGVQADVGSALNLKIPGVSLSTOICREFSLTL
 Y05204 len: 24 1 Human CLIP immunomodulatory peptide. Use of
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
 x(P)X(P)(P)XX(V)(S)(K)XX(W)XXX(L)x(7)
 KRKPPRPVSKRMATPFLMQLPPLM
 1:
 Y11993 len: 102 1 Human 5' EST secreted protein SEQ ID No: 53
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
 x(80)(L)X(T)(L)XX(W)(T)(H)XX(L)XXX(G)x(13)
 MDQFPESTVENFEYDDLAELACTIGDIVFGLVSLTSTVIAPIGLVGLLVVFLTSSKPKRS
 1:
 Y05545 len: 243 1 Wheat Type III glutathione transferase subu
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
 x(91)(L)X(S)(P)XX(V)(S)(R)XX(A)XXX(A)x(136)
 MAGGELIKLGGWAPGSPVYLRAQMLAVKGLSTYDLDPEDRWSISDILLIANSPPYKKVPLVING
 1:
 PMFEAAGRLAGLPIIDPAPFKLAAMERSVAEPKILLPGVRLKEEYITLALY
 PKNNIATGON

Y07041 len: 207 1 Breast cancer associated antigen precursor
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
 x(158)(A)X(S)(P)XX(I)(T)(H)XX(P)XXX(P)x(33)
 MSPILRLILALTLAPQAVSQPDAPGQRKRVSMIDVYTRATCQPREVVVPLVETLMGTAVX

AAPISTTSALTROPAAAAADAASVAKGCA

Y07007 len: 177 1 Breast cancer associated antigen precurs
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
 x(47)(V)X(S)(P)XX(I)(S)(K)XX(A)XXX(G)x(114)
 LCMHCKCKSCNNELTTRLEQNTLVNCTOPIADCPAOPGAFIRAIIVSSPEVYSIKYKALING
 1:
 W93953 len: 352 1 Human regulatory molecule HRM-9 protein.
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
 x(117)(P)X(S)(P)XX(L)(T)(K)XX(L)XXX(A)x(319)
 MHVAPASLALSGNINLSPSPCITKTLAPPAAPSLIANSOSGRHRAEAPLERDSGTSLIG
 1:
 GWRLEPEYTLISQEGPARKREYTTICRSLSEFELGKSGAKOAKRNAEKFLAFSN
 ISPENHISLTNNVGHSLGCTWHSLSRNSPEKINILKSLISIPRTDVLQLSIAEQGNITVYIDELSLANGOCYOLAEI
 Y04959 len: 273 1 Mycobacterium species protein sequence 4
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
 x(185)(P)X(T)(V)XX(W)(S)(H)XX(A)XXX(L)x(172)
 MANRGNAGGPPLESLDRDDHGHGHLARIGKRYLVRGVEILRTLLARAEVADVLELAP
 1:
 BALKVNARPLTVAEMSHLAGHGLVVEHVAVTASMLLOPRVADBSGLGNFRAG
 NLLIHRARRRVILMRTFRHRRERILAVAIYVAKRPHVDS
 Y04960 len: 280 1 Mycobacterium species protein sequence 4
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
 x(192)(P)X(T)(V)XX(W)(S)(H)XX(A)XXX(L)x(172)
 KPAEYGSANRGNAGGPPLESLDRDDHGHGHLARIGKRYLVRGVEILRTLLARAEVADVLELAP
 1:
 DLRQSLAARLAVNARPLTVAEMSHLAGHGLVVEHVAVTASMLLOPRVADBSGLGNFRAG
 GALRRAGNLLIHRARRRVILMRTFRHRRERILAVAIYVAKRPHVDS
 Y04901 len: 72 1 Mycobacterium species protein sequence 3
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
 x(50)(P)X(S)(I)XX(A)(T)(R)XX(A)XXX(S)x(6)
 ANRREGSATIALPAPVILNNTSOPRTLEPCGRINDPMSVTRITATTKALAPGSISRAIARRS
 1:
 Y04743 len: 110 1 Mycobacterium species protein sequence 1
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
 MAGQDELLEFDVPLVTLAEASRYLVERRALTLWADQERRANAPAVOGPIADFASV
 1:
 Y07239 len: 141 1 Fragment of human placental STAT3 protei
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
 x(70)(I)X(S)(V)XX(V)(T)(K)XX(L)XXX(S)x(155)
 WLDNIIDLVAKKIILALMNGEYIMGSLSKERRALTLSTKPGTFLRFSSESKGVTFTWEEK
 1:
 Y07240 len: 141 1 Fragment of human hepatic STAT3 protein
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
 x(70)(I)X(S)(V)XX(V)(T)(K)XX(L)XXX(S)x(155)
 WLDNIIDLVAKKIILALMNGEYIMGSLSKERRALTLSTKPGTFLRFSSESKGVTFTWEEK
 1:
 Y07241 len: 141 1 Fragment of mouse hepatic STAT3 protein
 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)
 x(70)(I)X(S)(V)XX(V)(T)(K)XX(L)XXX(S)x(155)
 WLDNIIDLVAKKIILALMNGEYIMGSLSKERRALTLSTKPGTFLRFSSESKGVTFTWEEK

Mon Feb 14 08:07:18 2000

ags.cat

Page 18

Databases searched:

GenSeq; Release 36.3, Released on 8Jul1999, Formatted on 9Jul1999

Total finds: 245
Total length: 23,686,106
Total sequences: 188,963
CPU time: 03:36.13

11AA_SEQUENCE 1.0
ID P90009 standard; protein: 102 AA.
AC P90009;
DE 1-NOV-1989 (first entry)
DT Mouse gamma-2b chain
KW Immunoglobulin E competitor; low affinity receptor sites;
KW Benrich enumeration; reduced hypersensitivity.
OS Murine
PN W08904834-A.
PD 01-JUN-1989.
PE 18-NOV-1988: G01018.
PR 19-NOV-1987: GB-027045.
PA (RECO) Research Corp Ltd.
PI Gould HU, Helm BA, Marsh PJBH;
DR WPI: 89-178361/24.
N-PSDB: N90237.
PT New immunoglobulin E competitor
PT - binding exclusively with low affinity receptors (FCER2)
PS Disclosure; Table 4; 33pp; English.
CC Amino acid residues (see corresp. DNA N90237) representing the
CC mouse gamma-2b chain. Used to extend P90006 (as Region.2) to
CC create an IGE competitor binding to type 2 low affinity
CC FC receptors
SQ Sequence 102 AA;
P90009 Length: 102 February 11, 2000 15:48 Type: P Check: 7203

1 POVYLIPPPA EOLSRKDVSL TCIIVGFNPG DISWEWTSNG HIEENYKDTA
51 PVLSDSGSYF IYSKLDIKTS KWERTDSFSC NVREHGLKNY YLKRTISRSP
101 GK
11AA_SEQUENCE 1.0
ID P91884 standard; peptide: 48 AA.
AC P91884;
DE 26-APR-1990 (first entry)
DT Antigenic Epstein-Barr virus peptide
KW Epstein-Barr virus; EBV; EBV antigen; EBV vaccine; EBV peptide.
PN EP-316170-A.
PD 17-MAY-1989.
PE 10-NOV-1988: 310591.
PR 11-NOV-1987: IE-003041.
PA (GGEN-) G Gene Galway Ltd.
PI Gannon BFX, Wallis DRG;
DR WPI: 89-146691/20.
N-PSDB: N91751.
PT Identifying antigens from a pathogenic organism expressed in vivo
PT - by generating a gene bank using nucleic acid
PT fragments and screening with infected serum
PS Claim 8(3); 21pp; English.
CC It is coded for by a DNA sequence isolated by using sera from patients
CC infected with EBV to screen colonies of micro-organisms transformed with
CC fragments of EBV DNA.
SQ Sequence 48 AA;
P91884 Length: 48 February 11, 2000 15:48 Type: P Check: 718

1 RGFPCSLCP SEEPETSGTP EPIGPASRRP POLRSPSPV KPRDCLRG
11AA_SEQUENCE 1.0
ID P94157 standard; protein: 246 AA.
AC P94157;
DT 12-JUN-1990 (first entry)

DE Acetyl-CoA reductase.
KW Polyester biopolymers; acetyl-CoA reductase; polyhydroxybutyrate; ss.
OS Alcaligenes eutrophus H16.
PN W08900202-A.
PD 12-JAN-1989.
PE 27-JUN-1988: U02227.
PR 29-JUN-1987: US-067695.
PA (MASI) Massachusetts Institute of Technology.
PI Peoples OP, Sinskey AJ;
DR WPI: 89-039655/05.
N-PSDB: N91209.
PT Constructing new polyester biopolymers - using genes encoding
PT beta-ketothiolase(s), acetoacetyl-CoA reductase(s) and
PT polyhydroxybutyrate synthetase(s).
PS Disclosure; P; English.
CC The enzyme is used to study the polyhydroxybutyrate (PHB) biosynthetic
CC pathway. It catalyses the reduction of acetoacetyl CoA to form
CC D(-)-beta-hydroxybutyryl-CoA, the substrate for PHB synthetase, and can
CC be used to control biopolymer syntheses and produce new biopolymers.
CC The sequence was deduced from the DNA sequence.
SQ See also P94154-P94156.
Sequence 246 AA;
P94157 Length: 246 February 11, 2000 15:48 Type: P Check: 2829

1 MTORIAVVTG GNGGIGTALC QRLAKDGRV VAGCGPNSPR REKWLDOOKA
51 LGFDFIASEG NVADWDSTRT AFDKXSEYV EYDVLINAG ITRDVERKM
101 TRADMDAVID TMLTSLFNT KQVIDGMADR GNGRIYNISS VNGOKGQFGQ
151 TNYSTAKAGL HGFMTALAQE VATKGVATNT VSPGRIATDM VKAIRQVLD
201 KIVATIPYKR LGLPEIASI CAMLSSESG FSTGADFSLN GGLMG
11AA_SEQUENCE 1.0
ID P90460 standard; protein: 288 AA.
AC P90460;
DE 1-NOV-1989 (first entry)
DT Alpha-factor profibrolase from Southern Copperhead venom
KW Southern copperhead fibrolase; snake venom fibrolase;
KW alpha-factor profibrolase fusion construct; Crotalidus;
KW thrombolytic agents; plasmin PK3308.
OS Actistodon confortrix confortrix
PN EP-323722-A.
PD 12-JUN-1989.
PE 16-DEC-1988: 311924.
PR 18-DEC-1987: US-134981.
PA (CHIR) Chiron Corp.
PI Valenzuela P, Brake A, Randolph A;
DR WPI: 89-200623/28.
N-PSDB: N90243.
PT DNA encoding snake venom fibrolase
PT - for use as thrombolytic agent.
PS Disclosure; fig 8; 41pp; English.
CC Alpha-factor profibrolase fusion construct (see N90243) of
CC plasmin PK3308. Used as a thrombolytic agent to treat
CC vertebrates, esp. humans. Produced by recombinant techniques.
SQ Sequence 288 AA;
P90460 Length: 288 February 11, 2000 15:48 Type: P Check: 560

1 MRPSIFTAV LPAASSALAA PVNTTDEET AOIPAEAVIG YLDEGDFDV
51 AYLFPNSTN NGLLFINNTI ASIAKKEGV SLDKROORFP QRYVOLIVA
101 DHRMNTKYNQ DSDKIRQWVH QIVNTIWEY RELNIOFTLY GLEIWSNDL
151 ITVTSVSHDT LASFGWRET DLRLRQHDN AOLLAIDBD GDYGLAYV
201 GMDCLKHSTG VIODHSAINL LVALTMAHEL GHNLCNHDG NOCHGANSNC
251 VMAALSDOP SKLFSDCKR DYOTFLTVNN POCILNKP

!!AA-SEQUENCE 1.0
ID P81999 standard; protein; 25 AA.
AC P81999;
DE 18-OCT-1990 (first entry)
DE Binding peptide 2 capable of binding to rotavirus capsid protein VP6
DE rotavirus strain S-All; VP6 inner capsid protein; binding peptide;
KW gastroenteritis; ss.
OS Rotavirus strain S-All.
PN EP-259149-A.
PD 09-MAR-1988.
PE 02-SEP-1987; 307746.
PR 03-SEP-1986; US-903222.
PA (UYSA-) Univ Saskatchewan.
PI Sabara M I, Frenchick P J, Mullin-Ready K F;
DR WPI: 88-065927/10.
DR Inducing immune response with complex of epitope contg molecule -
PT and rotavirus VP6 inner capsid protein as carrier, attached
PT without chemical coupling agent
PS Claim 9; Page 22; 30pp; English.
CC This peptide is found to bind to rotavirus protein VP6. It is
CC derived from rotavirus VP3 protein. Ability to bind is due to
CC the spatial arrangement of a Cys and an Arg in the 3-D protein.
CC VP6 is used as a carrier in a compsn for inducing immunological
CC response in mammals to a selected epitope. The carrier protein is
CC coupled to an epitope-bearing molecule via a binding sequence.
CC (See also P81997 and P81999-P82000).
CC Use of the VP6 inner capsid protein as carrier eliminates the need
CC for coupling agents (protein-protein interaction used) and avoids
CC adverse effects or antigenicity.
SQ Sequence 25 AA;
P81999 Length: 25 February 11, 2000 15:48 Type: P Check: 4995 ..

1 CNIPASIVS RNIVYTRAP NODIA

!!AA-SEQUENCE 1.0
ID P70477 standard; protein; 194 AA.
AC P70477;
DE 05-APR-1991 (first entry)
DE Sequence of human respiratory syncytial virus (HRSV) A2 strain
DE 22K protein.
KM Vaccine.
OS Human respiratory syncytial virus (HRSV).
PN MO8704185-A.
PD 16-JUL-1987.
PE 23-DEC-1986; 002756.
PR 14-JAN-1986; US-818740.
PA (UNIC-) UNIV OF N CAROLINA.
PA (WERT/) WERTZ G W.
PA (WERT/) WERTZ G W.
PA WPI: 87-206300/29.
DR N-PSDB: N70785.
DR Vaccines for human respiratory virus - comprising proteins or
PT fragment encoded by a DNA sequence coding for human respiratory
PT syncytial virus proteins.
PS Disclosure: Chart 14; 57pp; English.
CC A novel plasmid which comprises a DNA sequence encoding this
CC protein, and the protein itself, are claimed, for use as HRSV
CC vaccines. The vaccine can be administered to pregnant women or to
CC women of child bearing age to stimulate maternal antibodies.
CC Infants can also be vaccinated at 2-3 months of age.
SQ Sequence 194 AA;
P70477 Length: 194 February 11, 2000 15:48 Type: P Check: 9828 ..

1 MSRRNPCKFE IRGCHLNKR CHFSHNYFEW PPHALVYRON FMLNRLKSM
51 DKSIDTLEI SGAAELDRTE EYALGVYVL ESYIGSINNI TKOSACVAMS
101 KLTELNSDD IKRLDNEEL NSPKIRVYNT VISYIESNRK NNGQTHLKL
151 RLPADVLRKT IKNTLDIHS ITINPKEST VSDINDHAKN NDTT

!!AA-SEQUENCE 1.0
ID P70785 standard; protein; 194 AA.
AC P70785;
DE 05-APR-1991 (first entry)
DE Sequence encoding human respiratory syncytial virus (HRSV) A2 strain
DE 22K protein.
KM Vaccine; ss.
OS Human respiratory syncytial virus (HRSV).
PN MO8704185-A.
PD 16-JUL-1987.
PE 23-DEC-1986; 002756.
PR 14-JAN-1986; US-818740.
PA (UNIC-) UNIV OF N CAROLINA.
PA (WERT/) WERTZ G W.
PA (WERT/) WERTZ G W.
PA WPI: 87-206300/29.
DR N-PSDB: P70477.
DR Vaccines for human respiratory virus - comprising proteins or
PT fragment encoded by a DNA sequence coding for human respiratory
PT syncytial virus proteins.
PS Disclosure: Chart 14; 57pp; English.
CC A novel plasmid which comprises a DNA sequence encoding this
CC protein, and the protein itself, are claimed, for use as HRSV
CC vaccines. The vaccine can be administered to pregnant women or to
CC women of child bearing age to stimulate maternal antibodies.
CC Infants can also be vaccinated at 2-3 months of age.
SQ Sequence 194 AA;
P70785 Length: 194 February 11, 2000 15:48 Type: P Check: 9828 ..

1 MSRRNPCKFE IRGCHLNKR CHFSHNYFEW PPHALVYRON FMLNRLKSM
51 DKSIDTLEI SGAAELDRTE EYALGVYVL ESYIGSINNI TKOSACVAMS
101 KLTELNSDD IKRLDNEEL NSPKIRVYNT VISYIESNRK NNGQTHLKL
151 RLPADVLRKT IKNTLDIHS ITINPKEST VSDINDHAKN NDTT

!!AA-SEQUENCE 1.0
ID R24296 standard; protein; 384 AA.
AC R24296;
DE 20-NOV-1992 (first entry)
DE Regulatory protein Vans involved in glycopeptide resistance.
DE Glycopeptide antibiotic; Vancomycin; telcoplanin; resistant;
KW D-Ala-D-Ala ligase; peptidoglycan precursor; transposon;
KM Inverted repeats
OS Enterococcus faecium BM4147.
PN WO9207942-A.
PD 14-MAY-1992.
PE 29-OCT-1991; F00855.
PE 31-OCT-1990; FR-013579.
PA (INSP) INST PASTEUR.
PA Arthur M, Courvalin P, Duka-malen S, Molinas C;
PI WPI: 92-183677/22.
DR N-PSDB: Q25178.
DR Polyepitopes involved in expression of glycopeptide antibiotic
PT resistance - useful in diagnosing presence of Gram-positive
PT enterococcal strains e.g. Enterococcus faecium and E gallinarum
PS Claim 5; Fig 8; 163pp; French.
CC Vans is coded for by a 7.3kb HindIII-EcoRI fragment of
CC the plasmid p1816. (The plasmid contains resistance genes from
CC Enterococcus faecium BM 4147 and was described in New England J.
CC Med.; 319:157-161). The Vans protein is part of a two component
CC regulatory system for controlling the amount of expression of the
CC vancomycin-resistance proteins Vahr, Vahr and VahrX. The other
CC component of the regulatory system is the Vahr transcription
CC regulator. Vans protein can modify phosphorylation of Vahr in
CC response to vancomycin and thereby stimulates Vahr-dependent
CC transcription. See also Q25178-Q25183.
SQ Sequence 384 AA;
R24296 Length: 384 February 11, 2000 15:48 Type: P Check: 3600 ..

1 LVTKLKRNK DYSKLERLY MYVAIVVA IYFVLYIRSM IRGKLGDWIL
 51 SLENNYDAN HLDAMKLYQ SINNNIDIFI YVAIYISLI LCVNMLSKFA
 101 KYDEINTCI DVLIONERQ IELSAEMDA EOKLNTLKRT LERREDAKL
 151 AEORNDVVA YLADIKTPL TSIIGLSLI DEAPMPYDQ KARYHITLD
 201 KAYRELID EFFETRYNL QITTLKTHI DLYMLVOMT DEFYPOLSAH
 251 GKOAVIHAE DLTWSCDPK LARFNNIIL NAAAYSEDNS IIDITAGLSG
 301 DVYSIEFKMT GSIPDKLAA IFEKRYLDN ARSSDTGAG LGLAIKEII
 351 VHGGOIVAE SNDNTTFRV ELPAMPDLVD KRRS

11AA SEQUENCE 1.0
 ID R10995 standard; protein: 398 AA.

AC R10995;
 DT 13-MAY-1991 (first entry)
 DE Xenopus Bone Morphogenetic Factor BMP-2A.
 KW BMF; osteoporosis; fracture; cartilage.
 OS Xenopus laevis.
 PN EP-416578-A.
 PD 13-MAR-1991.
 PE 05-SEP-1990; 117079.
 PR 06-SEP-1989; JP-229250.
 PR 20-JUL-1990; JP-190774.
 PA (TAKE) TAKEIDA CHEMICAL IND KK.
 PA (SCIT-) SCITECH RESEARCH CO.
 PI Murakami K, Ikeno N, Kato Y;
 PI WPI: 91-075112/11.
 DR N-PSDB: Q10895.
 PT Xenopus laevis bone morphogenetic protein and DNA encoding it -
 used in therapy of fracture or osteoporosis
 PS Claim 2; Fig 4; 28pp; English.
 CC A xenopus laevis unfertilised egg cDNA library in lambda gt10 was
 screened with a PstI-HindIII fragment of X.laevis xar14 chromosomal
 CC DNA. Three clones were isolated, including clone Xbr22 which was
 found to encode a protein having homology with X.laevis BMP-2A. They
 CC were subcloned in pUC19 and used to transform competent E.coli HB101
 CC cells. Transformant E.coli HB101/pXbr22 coding for the BMP-2A was
 CC sequenced and the amino acid sequence of BMP-2A deduced from it.
 CC See also Q10890-4 and Q10896-7.
 SO Sequence 398 AA.

R10995 Length: 398 February 11, 2000 15:48 Type: P Check: 1620 ..

1 MWAGHISLL LQFYQILLG CTGLVDEEK RKYSESTRSS PQOSQOVLDO
 51 FELRLINMG LKRRPTGKN VVIPYMLD YHLSAQIAD DQGSSEVDYH
 101 MERAASRANT VRSFHHESM EEIPESGKT IQRFENLSS IPDEELVTSS
 151 ELRIPEQVQ EPIKTDGSKL HRINIYDIK PAAASRPV VRLDDRLLH
 201 HNSKKESTD VTPAITMIA HKOPNGEVV EYTHLDNTN VPKRHVRISK
 251 SLTDKGHP RIRPLVTF S HDKGHALK ROKROARHKO RKRLLKSCRR
 301 HPLVYDFSDV GWNMDIAPV GYHAFYCHG CPEPLADHLN STNHAIVQTL
 351 VNSVNTNIRK ACVPTELSA ISMLYLDENE KVLKNYODM VVEGCGCR

11AA SEQUENCE 1.0
 ID R24190 standard; protein: 186 AA.

AC R24190;
 DT 24-NOV-1992 (first entry)
 DE Bovine RSV strain A 51908 M2 protein.
 KW Bovine respiratory syncytial virus; vaccine; diagnosis; antibodies;
 KM M2 gene; BRSV.
 OS Bovine respiratory syncytial virus strain A 51908.

PN W09207940-A.
 PD 14-MAY-1992.
 PR 04-NOV-1991; U08177.
 PR 05-NOV-1990; US-608937.
 PA (SAMA/) SAMA S K.
 PI Samal SK.
 PI WPI: 92-183675/22.
 DR N-PSDB: Q25033.
 PT Bovine respiratory syncytial virus genes - used in the prodn. of
 agents for use in detection and as vaccines for BRSV infection.
 PS Claim 32; Page 57; 74pp; English.
 CC This is the sequence of bovine respiratory syncytial virus (BRSV)
 CC strain A51908 M2 protein. It can be used in the detection of BRSV
 CC antibodies and in vaccines to prevent infection. It can also be
 CC used for the production of BRSV protein antibodies.
 CC See also R24184-R24191 and R25310.
 SO Sequence 186 AA.

R24190 Length: 186 February 11, 2000 15:48 Type: P Check: 5872 ..

1 MSRRNPKRYE INGHCLNGK CHESHNTFEW APHALIVRON FMLNKILKSM
 51 DRNDTISEI SGAELDRTE EYALGVIGVL ESYLSSINNI TKQACVYAMS
 101 KLAEINND IRLRNKEVP TSPKIRIYNT VISYIDSNKR NTKOTIHLK
 151 RLPAVLKKT IKNTIDINE INGNNGDIN VDEONE

11AA SEQUENCE 1.0
 ID P81092 standard; protein: 95 AA.
 AC P81092;
 DT 21-MAR-1991 (first entry)
 DE Sequence of rhinovirus HRV2 viral protein P2B.
 KW Passive immunity; diagnosis; therapy; ss.
 OS Rhinovirus.
 PN EP-261403-A.
 PD 30-MAR-1988.
 PE 20-AUG-1987; 112104.
 PR 17-JAN-1987; DE-701301.
 PA (BOEH) Boehringer Ingelheim.
 PI Ducheher M, Skern T, Sommergruber W, Neubauer C, Grundler P, Blaas,
 PI Kuchler E, Friesl L, Zorn M;
 PI WPI: 88-085735/13.
 DR New DNA corresponding to viral RNA of rhino-virus HRV89 - useful for
 PT prodn. of polypeptide(s) for stimulating immune system against HRV
 PT 89.
 PS Example; Fig 20; 66pp; German.
 CC DNA molecules corresponding to all or part of the RNA of rhinovirus
 CC strain HRV89 (Fig 4, N81390) is claimed, esp. the portion encoding
 CC the viral proteins VP1-VP4, P2A-P2C, P3A-P3C. Also claimed are the
 CC polypeptides encoded by any of these DNA molecules. The polypeptides
 CC are used for stimulating a protective immune response and for
 CC blocking cellular receptors. Ab are useful for assay and for
 CC purification of the corresp. antigen, and can also be used for the
 CC therapeutic and diagnostic applications.
 SO Sequence 95 AA.

P81092 Length: 95 February 11, 2000 15:48 Type: P Check: 3060 ..

1 GVTYIMLG EAFNGEVD S VKEHIAINP VGNISKIIR WMLRIISAMV
 51 IIRNSSDPQ TILATLLIG CSGSPWFLK ERFCKWTQIN YIHKE

11AA SEQUENCE 1.0
 ID P81093 standard; protein: 323 AA.

AC P81093;
 DT 21-MAR-1991 (first entry)
 DE Sequence of rhinovirus HRV2 viral protein P2C.
 KW Passive immunity; diagnosis; therapy; ss.
 OS Rhinovirus.
 PN EP-261403-A.
 PD 30-MAR-1988.
 PF 20-AUG-1987; 112104.

17-JAN-1987; DE-701301.
 PA (BOEH) Boehringer Ingelheim.
 PI Duechler M, Skern T, Sommergruber W, Neubauer C, Grundler P, Blaas,
 PI Kuchler E, Frasel L, Zorn M;
 DR WPI: 88-06573/13.
 PT New DNA corresponding to viral RNA of rhino-virus HRV89 - useful for
 PT prodn. of polypeptide(s) for stimulating immune system against HRV
 PT 89.
 PS Example: fig 20; 66pp; German.
 CC DNA molecules corresponding to all or part of the RNA of rhinovirus
 CC strain HRV89 (fig 4, N81390) is claimed, esp. the portion encoding
 CC the viral proteins VP1-VP4, P2A-P2C, P3A-P3C. Also claimed are the
 CC polypeptides encoded by any of these DNA molecules. The polypeptides
 CC are used for stimulating a protective immune response and for
 CC blocking cellular receptors. Ab are useful for assay and for
 CC purification of the corresp. antigen, and can also be used for the
 CC therapeutic and diagnostic applications.
 SO Sequence 323 AA;
 P81093 Length: 323 February 11, 2000 15:48 Type: P Check: 5281 ..
 1 SDSMLAKETE ACNARGLEW IGNKISKEIE WNKSM,POAQ LKVKYLNELK
 51 KINLYEKVE SLRAYDMKTO EKIKKEIDTL HDLSKREPL YASEAKRIKT
 101 LYIKCDNIK OKKCEPVAI VHGPPGAGK SITTNELAM ITNDSIYSL
 151 PPDPEYFPGI EDOQSVIMD DIMONPAGD MTLFCOMVS VTPIPPVADL
 201 PKGKAPDSR FYLCSTNHSI LPPPTISLP AMNRFFELDI DIYHNEKD
 251 POGKLVAAA FRPCDVNRI GNARCCPVC GRAVSFEDRN SONKYSILAQV
 301 YNIMEEDRR RQVVDVMTA IFO
 !!AA_SEQUENCE 1.0
 ID R10974 standard; Protein: 246 AA.
 AC R10974;
 DT 17-APR-1991 (first entry)
 DE Acetoacetyl CoA reductase enzyme.
 KM Polymer biopolymers; polyhydroxybutyrate; polyhydroxy alkanate;
 KM beta-ketothiolase; acetoacetyl CoA reductase.
 OS Alkaligenes eutrophus.
 PN MO9100917-A.
 PD 24-JAN-1991.
 PF 10-JUL-1990; US-37815.
 PR 10-JUL-1989; US-37815.
 PA (MASI) MASSACHUSETTS INST TECH.
 PI Peoples OP, Sinsky AJ;
 DR N-PSDB: Q10501.
 PT Construction and modification of polyester bio-polymers - by
 PT introduction of poly-hydroxy-butyrate and -alkanoate genes into
 PT bacteria or plants
 PS Disclosure; fig 3; 64pp; English.
 CC This Alkaligenes eutrophus acetyl CoA reductase is an enzyme
 CC which is essential to the biosynthesis of polyhydroxyalkanoate
 CC (PHA). The gene encoding this is contained in plasmid clone
 CC pA33, downstream from the thiolase gene. The use of recombinant
 CC methods for producing such enzymes, required for polyester bio-
 CC polymer synthesis, allows for the control and modification of the
 CC synthesis process.
 CC See also Q10499-500 and Q10502-03.
 SO Sequence 246 AA;
 R10974 Length: 246 February 11, 2000 15:48 Type: P Check: 3836 ..
 1 MROKIVYNG GMGIGTAIC ORLARDGEFV VAGCGPNSPR REKWLEQOKA
 51 LGDFIFASBG NVADMSTKI AFDKXSEVG EVDVLINAG ITRDVFERKM
 101 TRADMAVID TNLTSIFNVT KOYIDMADR GMRIVNISS VNGOGGOGQO
 17-JAN-1987; DE-701301.
 PA (BOEH) Boehringer Ingelheim.
 PI Duechler M, Skern T, Sommergruber W, Neubauer C, Grundler P, Blaas,
 PI Kuchler E, Frasel L, Zorn M;
 DR WPI: 88-06573/13.
 PT New DNA corresponding to viral RNA of rhino-virus HRV89 - useful for
 PT prodn. of polypeptide(s) for stimulating immune system against HRV
 PT 89.
 PS Example: fig 20; 66pp; German.
 CC DNA molecules corresponding to all or part of the RNA of rhinovirus
 CC strain HRV89 (fig 4, N81390) is claimed, esp. the portion encoding
 CC the viral proteins VP1-VP4, P2A-P2C, P3A-P3C. Also claimed are the
 CC polypeptides encoded by any of these DNA molecules. The polypeptides
 CC are used for stimulating a protective immune response and for
 CC blocking cellular receptors. Ab are useful for assay and for
 CC purification of the corresp. antigen, and can also be used for the
 CC therapeutic and diagnostic applications.
 SO Sequence 323 AA;
 P81093 Length: 323 February 11, 2000 15:48 Type: P Check: 5281 ..
 1 SDSMLAKETE ACNARGLEW IGNKISKEIE WNKSM,POAQ LKVKYLNELK
 51 KINLYEKVE SLRAYDMKTO EKIKKEIDTL HDLSKREPL YASEAKRIKT
 101 LYIKCDNIK OKKCEPVAI VHGPPGAGK SITTNELAM ITNDSIYSL
 151 PPDPEYFPGI EDOQSVIMD DIMONPAGD MTLFCOMVS VTPIPPVADL
 201 PKGKAPDSR FYLCSTNHSI LPPPTISLP AMNRFFELDI DIYHNEKD
 251 POGKLVAAA FRPCDVNRI GNARCCPVC GRAVSFEDRN SONKYSILAQV
 301 YNIMEEDRR RQVVDVMTA IFO
 !!AA_SEQUENCE 1.0
 ID R10974 standard; Protein: 246 AA.
 AC R10974;
 DT 17-APR-1991 (first entry)
 DE Acetoacetyl CoA reductase enzyme.
 KM Polymer biopolymers; polyhydroxybutyrate; polyhydroxy alkanate;
 KM beta-ketothiolase; acetoacetyl CoA reductase.
 OS Alkaligenes eutrophus.
 PN MO9100917-A.
 PD 24-JAN-1991.
 PF 10-JUL-1990; US-37815.
 PR 10-JUL-1989; US-37815.
 PA (MASI) MASSACHUSETTS INST TECH.
 PI Peoples OP, Sinsky AJ;
 DR N-PSDB: Q10501.
 PT Construction and modification of polyester bio-polymers - by
 PT introduction of poly-hydroxy-butyrate and -alkanoate genes into
 PT bacteria or plants
 PS Disclosure; fig 3; 64pp; English.
 CC This Alkaligenes eutrophus acetyl CoA reductase is an enzyme
 CC which is essential to the biosynthesis of polyhydroxyalkanoate
 CC (PHA). The gene encoding this is contained in plasmid clone
 CC pA33, downstream from the thiolase gene. The use of recombinant
 CC methods for producing such enzymes, required for polyester bio-
 CC polymer synthesis, allows for the control and modification of the
 CC synthesis process.
 CC See also Q10499-500 and Q10502-03.
 SO Sequence 246 AA;
 R10974 Length: 246 February 11, 2000 15:48 Type: P Check: 3836 ..
 1 MROKIVYNG GMGIGTAIC ORLARDGEFV VAGCGPNSPR REKWLEQOKA
 51 LGDFIFASBG NVADMSTKI AFDKXSEVG EVDVLINAG ITRDVFERKM
 101 TRADMAVID TNLTSIFNVT KOYIDMADR GMRIVNISS VNGOGGOGQO
 !!AA_SEQUENCE 1.0
 ID R12468 standard; Protein: 248 AA.
 AC R12468;
 DT 29-AUG-1991 (first entry)
 DE Lufia cylindrica bioactive protein.
 KM Protein synthesis inhibition.
 OS Lufia cylindrica.
 PN J03109398-A.
 PD 09-MAY-1991.
 PF 22-SEP-1989; JP-247410.
 PR 22-SEP-1989; JP-247410.
 PA (SUMO) SUMITOMO CHEM IND KK.
 DR WPI: 91-181458/25.
 PT New protein of Lufia cylindrica, extracted from seeds - is used for
 PT inhibiting protein synthesis in viral proliferation.
 PS Claim 1; Fig 1; 6pp; Japanese.
 CC The protein is prepd. by extraction from seeds of L. cylindrica.
 CC It has protein synthesis inhibiting activity and is useful for
 CC inhibiting viral growth.
 CC See also J0112999.
 SO Sequence 248 AA;
 R12468 Length: 248 February 11, 2000 15:48 Type: P Check: 7673 ..
 1 DVRESLSSGS STSYSKEFGD LKRAKPSNGT VYNITILLSS ASGASRYTLM
 51 TLSNYDCKAI TVAVDVOLY IMGYLVNSTS YFENESDAKL ASQYFKSGT
 101 IVTLPYSGNY EKLQTRAGKI REKIPLEGPA LDSALITTFH YDSTAAAAF
 151 LVIIQTAAE SRFKYLEGOI IERISKNOVP STATISLENS LMSALSKOLO
 201 LAQTNGGTFK TPVYITDCKG QWEIITNVTG KVTAKNIOLL LNKONVA
 !!AA_SEQUENCE 1.0
 ID R13252 standard; Protein: 247 AA.
 AC R13252;
 DT 11-OCT-1991 (first entry)
 DE Murine Cytotoxic Cell Protease-1.
 KM Mouse; CCP-1 inhibitor; cytotoxic T-lymphocytes; ss.
 OS Mus musculus.
 PN W09110685-A.
 PD 25-JUL-1991.
 PF 17-JAN-1991; US-467880.
 PR 19-JAN-1990; US-467880.
 PA (SERA-) SERAGEN INC.
 PI Bleackley RC, Lode CG, Paetkau VH, James MN, Murphy M;
 DR WPI: 91-237989/32.
 PT DNA vectors, and inhibitors of cytotoxic cell protease - for
 PT treatment of auto-immune diseases e.g. pernicious anaemia,
 PT rheumatoid arthritis, allograft rejection etc.
 PS Claim 5; Fig 3; 62pp; English.
 CC The CCP-1 coding sequence was isolated from the cytotoxic T-cell
 CC line MTL 2.8.2 generated from GB8/J mice. The amino acid sequence
 CC of CCP-1 was predicted from the cDNA sequence. The structure of the
 CC protein was used to design peptides which competitively inhibit the
 CC protease. See also Q12863-6 and R13254-R13262.
 SO Sequence 247 AA;
 R13252 Length: 247 February 11, 2000 15:48 Type: P Check: 528 ..
 1 MKIILLTLT SLASRTAGE ITGCHVYKPH SRPYALALI KDOCPAALIG

51 GFLIREDFVL TNAHCEGSII NTLGAHNLIK EOEKTOQVIP MKXSIHPDPY
 101 NFKTFSNDIM LTKKSKAKR TRAVRPLNIP RRVNVKPGD VCVYAGWGMH
 151 AFIGKYSMTL QEVETLYQKD RCECESTFKNR YNKTNQICAG DPRKTRASFR
 201 GDSGGLPCK KVAAGIYSYG YKDGSPPPAF TVSSFLSWI KTKMSS

!!AA_SEQUENCE 1.0
 ID R20045 standard: protein; 292 AA.
 AC P60645.

DT 10-MAR-1993 (revised)
 DT 01-JAN-1980 (first entry)
 DE Mouse kidney cell Band 3-like anion transport protein.
 KW Band 3; anion transport protein; glycoprotein; cystic fibrosis;
 OS Mus musculus.
 PN M08605203.A.
 PD 12-SEP-1986.
 PR 06-MAR-1986; 000450.
 PR 07-MAR-1985; US-709183.
 PA (WHIT-) WHITEHEAD INST BIOM.
 PI Alper St, Kopito RR;
 DR WPI: 86-252231/38.
 DR NP-PDB: N60551.
 PT DNA coding for anion transport protein - esp. for detection of
 PS genetic defects
 CC Disclosure: Fig. 9; 49pp; English.
 CC Band 3 is the major glycoprotein of mammalian erythrocytes and
 CC functions as an anion transport in that it mediates the 1-for-1
 CC exchange of chloride (Cl-) and bicarbonate (HCO3-) across the
 CC erythrocyte plasma membrane. The anion transport protein may be
 CC used in the detection of genetic defects resulting in abnormal
 CC transport of ions across cell membranes, e.g. as in cystic fibrosis.
 CC This protein is specifically from mouse kidney cells. See also
 CC N60550 and P60644.
 SO Sequence 292 AA;

P60645 Length: 292 February 11, 2000 15:48 Type: P Check: 2381 ..

1 VDSIEDIYIT OKLSVPSGFS VTAPDKRWV INPLGEKTP PYMMNVASIL
 51 PAVVFLILF METQITLLI SKKEMLOKG SGHLLDLLI VAMGICALF
 101 GLPWIAATV RSVTHANALT VSKAVAPGD KPKIQEKQO RYTGILVALL
 151 VGLSMVIGDL LQGIPLAVLF GIFYMGVTS LNCIOFYERL HLLMPPKHQ
 201 PDVTVYKVR TMRMHLFTAL QLLCLALLMA VMSTASLAF PFILITVPL
 251 RMVVLRIET EREMKCLDAN EAEVPEDECE GVDEYNEMPM PV

!!AA_SEQUENCE 1.0
 ID R14247 standard: Protein; 149 AA.
 AC R14247;
 DT 03-JAN-1992 (first entry)
 DE HCV-T (6110-6557) encoded epitope.
 KW B10ptope; PCR diagnosis.
 OS Hepatitis C virus.
 PN M09114778.X.
 PD 03-OCT-1991.
 PR 28-MAR-1991; J00405.
 PR 28-MAR-1990; JP-080185.
 PR 13-JUN-1990; JP-154230.
 PR 14-JUN-1990; JP-153979.
 PR 09-NOV-1990; JP-305795.
 PA (MITSU) MITSU TOATSU CHEM INC.
 PI Takeda T, Enomoto N, Date T, Nakao T;
 DR WPI: 91-310579/42.
 DR NP-PDB: Q14076.
 PT New nucleotide sequences encoding HCV epitope(s) - for diagnosis
 PT of hepatitis C virus infection via polymerase chain reaction
 PS Disclosure: Fig 1(A)-(B); 79pp; Japanese.

CC The nucleotide sequences represented in Q14076-86 and Q14767-71
 CC encode epitopes from structural, non-structural and 5' untranslated
 CC domains of hepatitis C virus. The sequences are used for accurate
 CC and simple diagnosis and typing of HCV infection, using PCR
 CC amplification techniques.
 SO Sequence 149 AA;

R14247 Length: 149 February 11, 2000 15:48 Type: P Check: 3078 ..

1 VLKESKAKS TVKAKLSVE EACKLPPHS ARSKFGYGAQ DVRSLSRAI
 51 NHIRSYWMD LEDTEPIDT TIMANNEVFC VQPEKGRKP ARLIYFPDLG
 101 VVCEKEMALY DVVSTLPQAV MGSYGFQYS PGQRIEFLVN AMKSKCPM

!!AA_SEQUENCE 1.0
 ID R20004 standard: peptide; 25 AA.
 AC R20004;
 DT 25-MAR-1992 (first entry)
 DE VP6 inner capsid epitope; 84 TS, Peptide B.
 KW Rotavirus; vaccine; gastroenteritis; virus protein 6; hapten;
 KW immunogen; conjugate; carrier.
 OS Synthetic.
 PN US5071651-A.
 PD 10-DEC-1991.
 PR 05-MAR-1990; 489790.
 PR 03-SEP-1986; US-903222.
 PR 02-SEP-1987; US-092120.
 PR 05-MAR-1990; US-486750.
 PA (YUSA-) UNIV OF SASKATCHEWAN.
 PI Sabara M, Frenchick PJ, Mullin-Ready KF;
 DR WPI: 92-006804/01.
 PT New immunological carrier complex - contains VP6 polypeptide from
 PT rotavirus as the carrier mol.
 CS Claim 9b: Page 15; 23pp; English.
 CC The peptide binds to purified VP6 from bovine rotavirus strain C486.
 CC It was derived from the trypsin cleavage site of VP3 outer capsid
 CC protein and corresponds to amino acids 231-254. The cysteine at
 CC position 1 was added to facilitate coupling to a carrier protein.
 CC The peptide has an apparent mol. wt. of 2,734. The peptide can be
 CC used in vaccines against rotavirus, or to construct immunoconjugates
 CC with VP6 particles as carriers for haptens.
 CC See also R20003, and R21043-48, and R21071-73.
 SO Sequence 25 AA;

R20004 Length: 25 February 11, 2000 15:48 Type: P Check: 4995 ..

1 CNIPASIVS RNIVYTRAOP NODIA

!!AA_SEQUENCE 1.0
 ID R21409 standard: Protein; 482 AA.
 AC R21409;
 DT 01-APR-1992 (first entry)
 DE NADH dehydrogenase subunit 2.
 KW Pneumonia; assay; AIDS; immunosuppressed.
 OS Pneumocystis carinii.
 PN M09119005-A.
 PD 12-DEC-1991.
 PR 31-MAY-1991; G00869.
 PR 01-JUN-1990; GB-012196.
 PA (ISIS-) ISIS INNOVATION LTD.
 PI Wakefield AE, Hopkin JM, Moxon ER;
 DR WPI: 92-007487/01.
 DR NP-PDB: Q20065.
 PT New DNA sequences which act as oligo:nucleotide primers - for
 PT assaying DNA sample from respiratory secretion of a patient
 PT infected with P. carinii
 PS Claim 3: Fig 3; 42pp; English.
 CC The amino acid sequence is that of P. carinii NADH dehydrogenase
 CC subunit 2 which has translated from DNA from plasmid PA2112. See
 CC also R21410-R21413 and R20056.
 SO Sequence 482 AA;

R21409 Length: 482 February 11, 2000 15:48 Type: P Check: 617

1 MLSSITISOL IMAISSSHW NLVLSRISI ISLYSILIT YNYVVEIIG
 51 LGIGYNGFL QVTSIQFVD IFIFLLILIL LGITGYHYD KDNRELTSL
 101 YDROHLEFP VLSTFLRGS QILVSLNVI TFLSLLOS FSXVLSLR
 151 SSKOGLKFL LGALSCFLL LGFGLVYSYI GITSLESLAI FSKYNLNIYM
 201 QISLLICVLG ILEKIGIVP HONADIVYD VPIITITWLT TLTKISLLIF
 251 LMEFIYHSS ENMTITMLL SVLSYVCSL LGISOSRIR LLIYSMSHV
 301 GFLMLSISIM TEKSLEAFLE YLYOYSTNL NFFLIIMAG YFKNPDSO
 351 SPIIYNSLR GLVRVOPPLS ICLALISLST GGIPPIGFF GKNILYSTI
 401 TQGYLFISIL LVLASVLSIS YLEKVOULLF VGSLSLSPRN IQISTYLSFL
 451 IGVLTLMIAM FLVNPDEILO LINTICKYF IL

11AA-SEQUENCE 1.0
 ID R23004 standard; Protein: 297 AA.
 AC R23004;
 DT 29-OCT-1992 (first entry)
 DE Protein transcribed from the mda sequence of plasmid PKDSC50.
 KW Virulent; Salmonella Choleraesuis; mouse bacteremia; Salmonella.
 OS Salmonella choleraesuis.
 PN J04079885-A.
 PD 13-MAR-1992.
 PF 24-JUL-1990: JP-194069.
 PR (KIBU-) KIBUN KK.
 PA (KITA) KITASATO RES INST.
 DR WPI: 92-137918/17.
 DR N-PSDB: Q24046.
 PT New virulence plasmid PKDSC50 - provides DNA fragment useful as probe for detecting Salmonella.
 PS Disclosure; Fig 2, 9pp; Japanese.
 CC The 32 kd protein may be transcribed from the first reading frame of the "mda" fragment of virulent plasmid PKDSC50 from nucleotide 1159-2052. Three other proteins may also be transcribed from the Salmonella "mda" fragment, which is capable of causing mouse bacteremia.
 CC See also R23005-7.
 CC Sequence 297 AA;
 SQ

R23004 Length: 297 February 11, 2000 15:48 Type: P Check: 7059

1 MDPLINKRLK IFTLMETGS FSIATSVLYI TRPPLSRVIS DLRELKORL
 51 FIRKNGTILIP TEFQTIYRK VASHYIFLHA LEOEIGPTGK TKOLEIIFDE
 101 IYPSLKNLI ISALITISGOK TNIMRAVANS QIIEELCOTN NCIVISARNY
 151 FHRESIVCRT SVEGVALFI PKFPLCGKP DINLAGTPTV LFHEGAKNFN
 201 LDTIYHFEQ TLGITNPASF FPNVDFLSSL YRLQOGLAML LIPVRCAL
 251 GLSTDPHALHI KGVALCTSLX YPTKRREPD YRKAIKLIQ ELKOSTF

11AA-SEQUENCE 1.0
 ID R23695 standard; Protein: 286 AA.
 AC R23695;
 DT 10-NOV-1992 (first entry)
 DE blaC31 mutation (5)
 KW Beta-lactamase; BspMI excision linker; human growth hormone; antibiotic; transformants.
 OS Escherichia coli.
 PN key location/Qualifiers
 FT misc_difference 194..198
 FT /label=MC-16
 FT WO9207090-A.

PD 30-APR-1992. 007506.
 PF 21-OCT-1991; US-602158.
 PR 22-OCT-1990; US-602158.
 PA (GETH) GENENTECH INC.
 PI Botstein D, Palzkill T;
 DR WPI: 92-167168/20.
 DR N-PSDB: Q24133.
 PT Novel DNA modification method and prodn. of modified DNA
 PT libraries - used to determine effect of modification on
 PT interaction of encoded product, e.g. enzyme with target
 PS Disclosure; Fig 10; 98pp; English.
 CC The sequence given is the BspMI excision linker MC-16 which is
 CC one of a group of linkers designated blaC31.
 CC Within the scope of the invention, this sequence was used to
 CC illustrate a novel method of modifying DNA. DNA libraries
 CC expressing these modified polypeptides can be used to analyse the
 CC relationship between structure and function. Modified beta-lactamase
 CC libraries can be used to determine the susceptibility of a particular
 CC antibiotic to hydrolysis i.e. neutralization by mutant enzymes. Since
 CC all possible combinations of mutations were represented within the scope
 CC of the invention, a meaningful sample could be achieved. The method of
 CC the invention allows the measurement of successful transformants which
 CC contain a mutation linker, to determine the number of different colonies
 CC present in the library. This can then be used to calculate the
 CC probabilities that the most and the least common codon combinations are
 CC present in the library. The libraries can also be used to produce sets
 CC of modified polypeptides such as hormones, eg. human growth hormone, in
 CC order to detect and analyse structure/function relationships.
 CC Sequence 286 AA;
 SQ

R23695 Length: 286 February 11, 2000 15:48 Type: P Check: 4919

1 MSIOHFRVAL IPFAFCLP VFAHPETLVK VKDAEDOLGA RQVYIELDLN
 51 SKRISEFRP EERFPAMSTF KYLLGAVLS RVDAGORGLG RHVYSQNDP
 101 VERSPYTERH LTDMGTREL CSAITPMSD NTAAMLITTI GGHKEITAFI
 151 HNNGDVTRL DRCPELNEA IPNDERDTM PAAMATLRL LITASSVTLA
 201 SRQLDQWME ADKVAQPLR SALPAGWFLA DKSNGGERS RGIHALGPD
 251 GRSRTIVLY TTGSQAIMDE RNQQLAEIGA SLIHW

11AA-SEQUENCE 1.0
 ID R23601 standard; Protein: 85 AA.
 AC R23601;
 DT 18-JAN-1993 (first entry)
 DE NANBH virus antigen.
 KW Antibody; acute phase; chronic phase; vaccine.
 OS Non-A, non-B hepatitis virus.
 PN J04169193-A.
 PD 17-JUN-1992.
 PF 09-OCT-1990: 272980.
 PR 09-OCT-1990: JP-272980.
 PA (YAMA) YAMANOUCHI PHARM CO LTD.
 DR WPI: 92-253396/31.
 DR N-PSDB: Q26616.
 PT Antigenic polypeptide of non-A, non-B hepatitis virus - for
 PT diagnosing non-A hepatitis virus infections, also used as vaccine
 PS Claim 7; Page 2; 10pp; Japanese.
 CC The sequence given is an antigenic polypeptide of non-A, non-B
 CC hepatitis (NANBH) virus. This polypeptide is useful for the detection
 CC of NANBH virus antibodies. This antigen can detect the virus
 CC antibodies in both the acute phase and chronic phase. The detection of
 CC NANBH virus genome using this antigen is very sensitive. Polyclonal
 CC and monoclonal antibodies for the detection of the virus antigen and
 CC the medical treatment of NANBH, can be prepared using the antigenic
 CC peptide. This peptide can be used in the production of a vaccine.
 CC Sequence 85 AA;
 SQ

R23601 Length: 85 February 11, 2000 15:48 Type: P Check: 5700

1 KRYTEFLDLOV LDDHYRDVLEK EMKAKASTVK AKLLSVEAC KLTPPHSAKS
51 KFGYGAKDVR NLSSKAVNHI RSVKDLLED TETPL
11AA-SEQUENCE 1.0
ID R25869 standard; Protein: 239 AA.
AC R25869;
DE 21-JAN-1993 (first entry)
DE HCV polypeptide 16.
KM Recombinant vector: E. coli; diagnostic; reagent; type C hepatitis.
OS Hepatitis C virus.
PN J04179482-A.
PD 26-JUN-1992.
PE 11-NOV-1990: JP-304417.
PR 11-NOV-1990: JP-304417.
PA (TOKU) TOKUYAMA SODA KK.
DR WPI: 92-26363/32.
DR N-PSDB: Q26996.
PT Hepatitis C virus antigen expressed as recombinant in E.coli -
PT useful for diagnosis of hepatitis C virus infection
PS Claim 1; Page 5; 66pp; Japanese.
CC The sequences given in R25854-74 are hepatitis C virus proteins.
CC The genes encoding these proteins can each be used to prepare
CC recombinant vectors by ligating the gene of interest in to a vector
CC to be expressed in E. coli. These polypeptides are useful as
CC diagnostic reagents for type C hepatitis and they may be produced
CC efficiently by recombinant methods.
SQ Sequence 239 AA;
R25869 Length: 239 February 11, 2000 15:48 Type: P Check: 3783 ..
1 VCCSMSTWT GALITPCAE ESKLPINPLS NSLLRHSMV YSTERSASL
51 ROKVTEDRL QVDDHYRDV LKEMAKAST VKARLSIEE ACKLTPPHSA
101 KSKRGYAKD VRLSSRAVN HIRSWEDL EDETPIDIT IMAKREVCY
151 OPEKGGKRA RLIVFPDLOV RYCEKALYD VSTLPOAVM GSYGFOYSP
201 GQRYEPLVNT WMSKCPMGF SYDTRCPDST VTENDIRTE
11AA-SEQUENCE 1.0
ID R25891 standard; Protein: 253 AA.
AC R25891;
DE 21-JAN-1993 (first entry)
DE HK16.
KM Recombinant vector: E. coli; diagnostic; reagent; type C hepatitis.
OS Hepatitis C virus.
FH Key Location/Qualifiers
FT 13..251
FT /tag- a
FT /note- "Sequence R25869"
PN J04179482-A.
PD 26-JUN-1992.
PE 11-NOV-1990: 304417.
PR 11-NOV-1990: JP-304417.
PA (TOKU) TOKUYAMA SODA KK.
DR WPI: 92-26363/32.
DR N-PSDB: Q27018.
PT Hepatitis C virus antigen expressed as recombinant in E.coli -
PT useful for diagnosis of hepatitis C virus infection
PS Disclosure: Fig 17; 66pp; Japanese.
CC The sequences given in R25876-95 are encoded by the claimed hepatitis
CC C virus genes of the invention which have been inserted into an E.
CC coli vector. These polypeptides are useful as diagnostic reagents
CC for type C hepatitis and they may be produced efficiently by
CC recombinant DNA techniques.
SQ Sequence 253 AA;
R25891 Length: 253 February 11, 2000 15:48 Type: P Check: 4644 ..
1 MITSUBISHI QYVCGSMST WTGALITPCA AESKLPINP LSNLRRHS

51 MYSTSRSA SROKATFD FLOYDDHR DYKEMAKA STYKARLLSI
101 EEAOKLTPPH SAKSKGYGA KDVSLSRA VNRISWED LLEDETPID
151 TTMARNEVE CVOPKGRK PARLIVPDL GVRCEKMA YDVSTLPQA
201 VMGSPYGFQY SPGRVREFLV NTKSKKCPM GSYDTRCFD STVENDIRT
251 EGL
11AA-SEQUENCE 1.0
ID R25303 standard; Protein: 194 AA.
AC R25303;
DE 03-MAR-1993 (first entry)
DE HRSV 22K protein
KM Vaccine: human respiratory syncytial virus; HRSV; F; G; 22K; 9.5K;
KM major capsid protein; N
OS human respiratory syncytial virus strain A2.
PN US5149650-A.
PD 22-SEP-1992.
PE 14-JAN-1986: 818740.
PR 14-JAN-1986: US-818740.
PR 13-JUL-1988: US-218737.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Collins PL, Wertz GW;
DR WPI: 92-340247/41.
DR N-PSDB: Q29624.
PT Vaccines for human respiratory virus - include structural genes
PT coding for native structural viral proteins and immunogenic
PT fragments
PS Disclosure: Page 19; 21pp; English.
CC The sequences of mRNA encoding HRSV structural proteins are given in
CC CC 029622-26. The sequences are F, G, 22K, 9.5K and major capsid
CC protein N. The sequences and encoded proteins are useful for
CC preparing vaccines against HRSV. The vaccines can be used to confer
CC immunity against respiratory tract infections on human subjects.
SQ Sequence 194 AA;
R25303 Length: 194 February 11, 2000 15:48 Type: P Check: 9668 ..
1 MSRTNPKKEE IRGCLNGKR CHESHNYEWM PHALLVRON FMLNRIKSM
51 DKSIDTISEI SGAEELDRTE EYALGVVGL ESYIGSINI TKOSACVAIS
101 KLITELNSDD IKLRDNEEL NSPKIRYNT VISYIESNRK NNOITHLK
151 RLPAVDLAKT IKNTDIHKS ITINPKEST VSDTNDHAKN NDTT
11AA-SEQUENCE 1.0
ID R29874 standard; Protein: 171 AA.
AC R29874;
DE 26-APR-1993 (first entry)
DE HCV NS4-NS5 peptide N29-1, N29-2, N29-3.
KM Clone: polypeptide; NS4-NS5; Hepatitis C; Virus; HCV; serum; HC;
KM transcritpase; cDNA; primer; allele; core; region; upstream;
KM hydrophilic; turn structure; alpha helix; beta sheet; antigen;
KM determinant; antiserum.
OS Hepatitis C virus
FH Key Location/Qualifiers
FT misc_difference 61
FT /label- Glu, Lys
FT /label- Ala, Ser
FT misc_difference 155
FT EP-518313-A.
PD 16-DEC-1992.
PE 11-JUN-1992: 109812.
PR 11-JUN-1991: JP-139268.
PR 12-JUL-1991: JP-172794.
PR 07-OCT-1991: JP-287008.
PR 16-DEC-1991: JP-332329.
PR 20-APR-1992: JP-099957.
PA (MITU) MITSUBISHI KASEI CORP.

PI Hayashi N, Honda Y, Murakami T, Seki M, Takahashi K;
PI Teranishi Y;
DR WPI: 92-417213/51.
DR N-PSDB: 032505.
PI New hepatitis C virus gene and its encoded protein - used for
PI diagnosing and vaccinating against hepatitis C virus infections
PS Disclosure: Page 197-98: 305pp; English.
CC The sequences given in R29871-906 and R29533 are encoded by various
CC clones of the NS4-NS5 regions of the Hepatitis C virus (HCV) gene of
CC the invention. These NS4-NS5 RNA sequences were isolated from the
CC serum of a patient suffering from hepatitis C (HC). The isolated RNA
CC sequences were converted into cDNA using transcriptase in the presence
CC of one of the primer sequences given in 032565-77. The sequences were
CC then amplified using primer pairs. The cDNA sequences isolated
CC represent different alleles of the same region of the HCV gene.
CC Sequence analysis shows that these clones represent the core region
CC and some upstream sequences of HCV. These polypeptides are thought to
CC contain a highly hydrophilic region which can adopt a "turn structure"
CC which is not an alpha helix or a beta sheet. These polypeptides are
CC thought to act as antigen determinants and are highly reactive with
CC antiserum raised against HCV-associated antigens. See also Q32436.
SQ Sequence 171 AA;
R29874 Length: 171 February 11, 2000 15:48 Type: P Check: 9846 ..
1 YRDVLEKMK KASTVAKKL SVEACKLTP PHSARSKFGY GAKDVSLS
51 KAVNHRSW KDLEDTETP IDTTIMAKNE VFCVQPEKGG RKPARIIVFP
101 DLGVRCEKM ALYDVSTLP QAVMGSSYGF QSPGQRFVEF LVNAWSKRS
151 PMGFYDTRC FDSVTENDI R
11AA_SEQUENCE 1.0
ID R29879 standard; Protein: 285 AA.
DR R29879;
DE 26-APR-1993 (first entry)
KW HCV NS4-NS5 peptide; NS4-NS5; Hepatitis C; Virus; HCV; serum; HC;
KW Clone; polypeptide; cDNA; primer; allele; core; region; upstream;
KW hydrophilic; turn structure; alpha helix; beta sheet; antigen;
KW determinant; antiserum.
OS Hepatitis C virus
PN EP-518313-A.
PF 16-DEC-1992: 109812.
PD 11-JUN-1992: 109812.
PR 11-JUN-1991: JP-139268.
PR 12-JUL-1991: JP-172794.
PR 07-OCT-1991: JP-287008.
PR 16-DEC-1991: JP-332329.
PR 20-APR-1992: JP-099957.
PA (MRTU) MITSUBISHI KASEI CORP.
PI Hayashi N, Honda Y, Murakami T, Seki M, Takahashi K;
PI Teranishi Y;
DR WPI: 92-417213/51.
DR N-PSDB: 032510.
PI New hepatitis C virus gene and its encoded protein - used for
PI diagnosing and vaccinating against hepatitis C virus infections
PS Disclosure: Page 207-08: 305pp; English.
CC The sequences given in R29871-906 and R29533 are encoded by various
CC clones of the NS4-NS5 regions of the Hepatitis C virus (HCV) gene of
CC the invention. These NS4-NS5 RNA sequences were isolated from the
CC serum of a patient suffering from hepatitis C (HC). The isolated RNA
CC sequences were converted into cDNA using transcriptase in the presence
CC of one of the primer sequences given in 032565-77. The sequences were
CC then amplified using primer pairs. The cDNA sequences isolated
CC represent different alleles of the same region of the HCV gene.
CC Sequence analysis shows that these clones represent the core region
CC and some upstream sequences of HCV. These polypeptides are thought to
CC contain a highly hydrophilic region which can adopt a "turn structure"
CC which is not an alpha helix or a beta sheet. These polypeptides are
CC thought to act as antigen determinants and are highly reactive with
CC antiserum raised against HCV-associated antigens. See also Q32436.

SQ Sequence 285 AA;
R29879 Length: 285 February 11, 2000 15:48 Type: P Check: 8945 ..
1 YRDVLEKMK KASTVAKKL SVEACKLTP PHSARSKFGY GAKDVSLS
51 KAVNHRSW KDLEDTETP IDTTIMAKNE VFCVQPEKGG RKPARIIVFP
101 DLGVRCEKM ALYDVSTLP QAVMGSSYGF QSPGQRFVEF LVNAWSKRS
151 PMGFYDTRC FDSVTENDI R
201 GPLTNSKGN CGYRCRAG VLTTCGNTL TCYLRKSNAC RAKLDDCTM
251 LVCGDDLYVI CESAGTQEDA AMLRFTTAM TRNSA
11AA_SEQUENCE 1.0
ID R29895 standard; Protein: 171 AA.
DR R29895;
DE 26-APR-1993 (first entry)
KW HCV NS4-NS5 peptide; NS4-NS5; Hepatitis C; Virus; HCV; serum; HC;
KW Clone; polypeptide; cDNA; primer; allele; core; region; upstream;
KW hydrophilic; turn structure; alpha helix; beta sheet; antigen;
KW determinant; antiserum.
OS Hepatitis C virus
PN EP-518313-A.
PD 16-DEC-1992: 109812.
PF 11-JUN-1992: 109812.
PR 11-JUN-1991: JP-139268.
PR 12-JUL-1991: JP-172794.
PR 07-OCT-1991: JP-287008.
PR 16-DEC-1991: JP-332329.
PR 20-APR-1992: JP-099957.
PA (MRTU) MITSUBISHI KASEI CORP.
PI Hayashi N, Honda Y, Murakami T, Seki M, Takahashi K;
PI Teranishi Y;
DR WPI: 92-417213/51.
DR N-PSDB: 032526.
PI New hepatitis C virus gene and its encoded protein - used for
PI diagnosing and vaccinating against hepatitis C virus infections
PS Disclosure: Page 242-43: 305pp; English.
CC The sequences given in R29871-906 and R29533 are encoded by various
CC clones of the NS4-NS5 regions of the Hepatitis C virus (HCV) gene of
CC the invention. These NS4-NS5 RNA sequences were isolated from the
CC serum of a patient suffering from hepatitis C (HC). The isolated RNA
CC sequences were converted into cDNA using transcriptase in the presence
CC of one of the primer sequences given in 032565-77. The sequences were
CC then amplified using primer pairs. The cDNA sequences isolated
CC represent different alleles of the same region of the HCV gene.
CC Sequence analysis shows that these clones represent the core region
CC and some upstream sequences of HCV. These polypeptides are thought to
CC contain a highly hydrophilic region which can adopt a "turn structure"
CC which is not an alpha helix or a beta sheet. These polypeptides are
CC thought to act as antigen determinants and are highly reactive with
CC antiserum raised against HCV-associated antigens. See also Q32436.
SQ Sequence 171 AA;
R29895 Length: 171 February 11, 2000 15:48 Type: P Check: 8862 ..
1 YRDVLEKMK KASTVAKKL SVEACKLTP PHSARSKFGY GAKDVSLS
51 KAVNHRSW KDLEDTETP IDTTIMAKNE VFCVQPEKGG RKPARIIVFP
101 DLGVRCEKM ALYDVSTLP QAVMGSSYGF QSPGQRFVEF LVNAWSKRS
151 PMGFYDTRC FDSVTENDI R
11AA_SEQUENCE 1.0
ID R29886 standard; Protein: 171 AA.
DR R29886;
DE 26-APR-1993 (first entry)
KW HCV NS4-NS5 peptide; NS4-NS5; Hepatitis C; Virus; HCV; serum; HC;
KW Clone; polypeptide; cDNA; primer; allele; core; region; upstream;
KW hydrophilic; turn structure; alpha helix; beta sheet; antigen;
KW determinant; antiserum.
OS Hepatitis C virus
PN EP-518313-A.
PD 16-DEC-1992: 109812.
PF 11-JUN-1992: 109812.
PR 11-JUN-1991: JP-139268.
PR 12-JUL-1991: JP-172794.
PR 07-OCT-1991: JP-287008.
PR 16-DEC-1991: JP-332329.
PR 20-APR-1992: JP-099957.
PA (MRTU) MITSUBISHI KASEI CORP.
PI Hayashi N, Honda Y, Murakami T, Seki M, Takahashi K;
PI Teranishi Y;
DR WPI: 92-417213/51.
DR N-PSDB: 032526.
PI New hepatitis C virus gene and its encoded protein - used for
PI diagnosing and vaccinating against hepatitis C virus infections
PS Disclosure: Page 242-43: 305pp; English.
CC The sequences given in R29871-906 and R29533 are encoded by various
CC clones of the NS4-NS5 regions of the Hepatitis C virus (HCV) gene of
CC the invention. These NS4-NS5 RNA sequences were isolated from the
CC serum of a patient suffering from hepatitis C (HC). The isolated RNA
CC sequences were converted into cDNA using transcriptase in the presence
CC of one of the primer sequences given in 032565-77. The sequences were
CC then amplified using primer pairs. The cDNA sequences isolated
CC represent different alleles of the same region of the HCV gene.
CC Sequence analysis shows that these clones represent the core region
CC and some upstream sequences of HCV. These polypeptides are thought to
CC contain a highly hydrophilic region which can adopt a "turn structure"
CC which is not an alpha helix or a beta sheet. These polypeptides are
CC thought to act as antigen determinants and are highly reactive with
CC antiserum raised against HCV-associated antigens. See also Q32436.

KW Clone: polypeptide: NS4-NS5; Hepatitis C virus: HCV; serum: HC;
 KW transcriptase; cDNA; primer: allele; core; region: upstream;
 KW hydrophilic; turn structure: alpha helix; beta sheet; antigen;
 OS Hepatitis C virus
 PN EP-518313-A.
 PD 16-DEC-1992.
 PF 11-JUN-1992; JP-139268.
 PR 11-JUN-1991; JP-139268.
 PR 12-JUL-1991; JP-172794.
 PR 07-OCT-1991; JP-287008.
 PR 16-DEC-1991; JP-332329.
 PR 20-APR-1992; JP-099957.
 PA (MITU) MITSUBISHI KASEI CORP.
 PI Hayashi N, Honda Y, Murakami T, Seki M, Takahashi K;
 PI Teranishi Y;
 PI WPI: 92-417213/51.
 DR N-PSDB: Q32527.
 PT New hepatitis C virus gene and its encoded protein - used for
 PT diagnosing and vaccinating against hepatitis C virus infections
 PS Disclosure; Page 243-44; 305pp; English.
 CC The sequences given in R29871-906 and R29533 are encoded by various
 CC clones of the NS4-NS5 regions of the Hepatitis C virus (HCV) gene of
 CC the invention. These NS4-NS5 RNA sequences were isolated from the
 CC serum of a patient suffering from hepatitis C (HC). The isolated RNA
 CC sequences were converted into cDNA using transcriptase in the presence
 CC of one of the primer sequences given in Q32565-77. The sequences were
 CC then amplified using primer pairs. The cDNA sequences isolated
 CC represent different alleles of the same region of the HCV gene.
 CC Sequence analysis shows that these clones represent the core region
 CC and some upstream sequences of HCV. These polypeptides are thought to
 CC contain a highly hydrophilic region which can adopt a "turn structure"
 CC which is not an alpha helix or a beta sheet. These polypeptides are
 CC thought to act as antigen determinants and are highly reactive with
 CC antiserum raised against HCV-associated antigens. See also Q32436.
 SQ Sequence 171 AA;
 R29896 Length: 171 February 11, 2000 15:48 Type: P Check: 9862 ..
 1 YRDVLKEMKA KASTYKAKL SVEACKLTP PHSRASKRGY GANDVRSLS
 51 KAVNHRSW KDLEDTETP IDTTIMAKNE VFCVQPEKGG RKPARLIYFP
 101 DLGVRCEKM ALYDVSTLP QAVMGSSYGF QYSPQQRVEE LVNANKSKKS
 151 PMGFAYDTRC FDSYVTENDI R
 !!AA-SEQUENCE 1.0
 ID R29897 standard; Protein: 167 AA.
 AC R29897;
 DT 26-APR-1993 (first entry)
 DE HCV NS4-NS5 peptide N29-3.
 KW Clone: polypeptide: NS4-NS5; Hepatitis C virus: HCV; serum: HC;
 KW transcriptase; cDNA; primer: allele; core; region: upstream;
 KW hydrophilic; turn structure: alpha helix; beta sheet; antigen;
 OS Hepatitis C virus
 PN EP-518313-A.
 PD 16-DEC-1992.
 PF 11-JUN-1992; JP-139268.
 PR 11-JUN-1991; JP-139268.
 PR 12-JUL-1991; JP-172794.
 PR 07-OCT-1991; JP-287008.
 PR 16-DEC-1991; JP-332329.
 PR 20-APR-1992; JP-099957.
 PA (MITU) MITSUBISHI KASEI CORP.
 PI Hayashi N, Honda Y, Murakami T, Seki M, Takahashi K;
 PI Teranishi Y;
 PI WPI: 92-417213/51.
 DR N-PSDB: Q32528.
 PT New hepatitis C virus gene and its encoded protein - used for
 PT diagnosing and vaccinating against hepatitis C virus infections
 PS Disclosure; Page 244-45; 305pp; English.

CC The sequences given in R29871-906 and R29533 are encoded by various
 CC clones of the NS4-NS5 regions of the Hepatitis C virus (HCV) gene of
 CC the invention. These NS4-NS5 RNA sequences were isolated from the
 CC serum of a patient suffering from hepatitis C (HC). The isolated RNA
 CC sequences were converted into cDNA using transcriptase in the presence
 CC of one of the primer sequences given in Q32565-77. The sequences were
 CC then amplified using primer pairs. The cDNA sequences isolated
 CC represent different alleles of the same region of the HCV gene.
 CC Sequence analysis shows that these clones represent the core region
 CC and some upstream sequences of HCV. These polypeptides are thought to
 CC contain a highly hydrophilic region which can adopt a "turn structure"
 CC which is not an alpha helix or a beta sheet. These polypeptides are
 CC thought to act as antigen determinants and are highly reactive with
 CC antiserum raised against HCV-associated antigens. See also Q32436.
 SQ Sequence 167 AA;
 R29897 Length: 167 February 11, 2000 15:48 Type: P Check: 2886 ..
 1 YRDVLKEMKA KASTYKAKL SVEACKLTP PHSRASKRGY GANDVRSLS
 51 KAVNHRSW KDLEDTETP IDTTIMAKNE VFCVQPEKGG RKPARLIYFP
 101 DLGVRCEKM ALYDVSTLP QAVMGSSYGF QYSPQQRVEE LVNANKSKKS
 151 PMGFAYDTRC FDSYVTE
 !!AA-SEQUENCE 1.0
 ID R29909 standard; Protein: 277 AA.
 AC R29909;
 DT 05-MAY-1993 (first entry)
 DE Prod. of the luffin-f gene.
 KW Luffin; amplification; missile treatment; virus; agriculture.
 OS Luffin cylindrica.
 PN J0430687-A.
 PD 11-NOV-1992.
 PF 18-APR-1991; 086857.
 PR 18-APR-1991; JP-086857.
 PA (NISR) JAPAN TOBACCO INC.
 DR WPI: 92-426681/52.
 DR N-PSDB: Q32687.
 PT New gene coding luffin - for missile treatment and prevention of
 PT viral diseases in agriculture
 PS Disclosure; Page 10; 19pp; Japanese.
 CC The sequence is that of the prod. of the luffin-f gene of Luffa
 CC cylindrica. The gene sequence was obtained by PCR of luffa cDNA
 CC using a PCR primer designed based on the luffin protein sequence
 CC from amino acid residues 54-60. The luffin gene can be used in
 CC missile treatment and for prevention of viral diseases in agriculture.
 CC Luffin may be produced stably, by transforming a microbe such as E. coli
 CC with the luffin gene.
 CC See also R22910.
 SQ Sequence 277 AA;
 R29909 Length: 277 February 11, 2000 15:48 Type: P Check: 429 ..
 1 MKRFYLLILA IFVAASYEA DVAFSLSGSS STYSKFLGD LRKALSNGT
 51 VNYITLLLS ASGASRYTLM TLSNYDKAI TVADVNTWY IMGYLVNSIS
 101 YFPNESDARL ASQYVFKGST IYTLPRSGN EKLQTAAGKI REKIPLGFP
 151 LDSAITLFLH YDSTAAAMAF LVIIQTAAE SRFKIEGQI IERISNOVP
 201 SLATISLENE WSALSKQIQL AQTNNGTFKT PVTITDKGQ RVEITWTSK
 251 VTKNIQLUL NYKONVAAP EDVSAKH
 !!AA-SEQUENCE 1.0
 ID R29910 standard; Protein: 278 AA.
 AC R29910;
 DT 05-MAY-1993 (first entry)
 DE Prod. of the luffin-g gene.
 KW Luffin; amplification; missile treatment; virus; agriculture.

OS Luffa cylindrica.
 PN J04320687-A.
 PD 11-NOV-1992.
 PF 18-APR-1991; 086857.
 PR 18-APR-1991; JP-086857.
 PA (NIBB.) JAPAN TOBACCO INC.
 WP1: 92-42688/52.
 DR N-PSDB: Q32688.
 PT New gene coding luffin - for missile treatment and prevention of
 PI viral diseases in agriculture.
 PS Disclosure; Page 12; 12pp Japanese.
 CC The sequence is that of the prod. of the luffin-g gene of luffa
 CC cylindrica. The gene sequence was obtd. by PCR of luffa cDNA
 CC using a PCR primer designed based on the luffin protein sequence
 CC from amino acid residues 54-60. The luffin gene can be used in
 CC missile treatment and for prevention of viral diseases in agriculture.
 CC Luffin may be produced stably, by transforming a microbe such as E. coli
 CC with the luffin gene.
 CC See also R22909.
 SO Sequence 278 AA.

R29910 Length: 278 February 11, 2000 15:48 Type: P Check: 9157

1 MNRETFSL ILIAFTVE GANVSFSLG ADSKSYKFI TALRKALPSK
 51 EKVENIFLL PSAGSASRYI LMOISNDAR AITMAIDVTN VYMGIVNS
 101 TSYFENSDA KLASQVFKG STVTLFYSQ NYERLONAG KVEKPIPLGF
 151 RAFPASATSL FHYDSTAAAG AFLVIOTTA EASREKYLEG QIERIPKNE
 201 VPSFALSL EWSALSOKI QLAQTNGAF RPPVYIDNK GQRYEIKDVN
 251 SKVTNNIKL LNKONIAF DDGIPTKH

11AA SEQUENCE 1.0
 ID R31608 standard; Protein: 205 AA.
 AC R31608-1993 (first entry)
 DT Homologous to chicken nov gene exon 3-4-encoded protein.
 DE Homologous to chicken nov gene exon 3-4-encoded protein.
 KM avian nephroblastoma; avian myeloblastoma virus;
 KW stringent hybridisation; ss.
 PN WO9300430-A.
 PD 07-JAN-1993.
 PF 25-JUN-1992; F00589.
 PR 25-JUN-1991; FR-007807.
 PA (CNRS) CENT NAT RECH SCI.
 PI Martinerie C, Perbal B;
 DR WP1: 93-03637/04.
 PT Nucleotide sequences hybridising to regions of chicken nov gene -
 PT useful as probes for detecting complementary sequences to
 PT evaluate development and/or differentiation of tumours
 PS Claim 21; Page 39; 67pp; French.
 CC The chicken nov gene is stimulated in avian nephroblastoma induced
 CC by avian myeloblastoma virus but not in normal adult kidney. A
 CC 1975bp cDNA sequence (Q36031) was isolated from a gene bank prepared
 CC from chicken embryonic fibroblasts screened with a tumour-derived
 CC probe. Fragment XXII (Q36044) is part of the 3rd and 4th exons of
 CC the nov gene; nucleotide sequences which hybridise to Fragment XXII
 CC under stringent conditions (i.e. 50% formamide, 5 x SSC) are claimed.
 CC The claimed sequences preferably encode a protein with the sequence
 CC XXII (R31608) or an amino acid sequence 60% homologous to it.
 SO Sequence 205 AA.

R31608 Length: 205 February 11, 2000 15:48 Type: P Check: 7012

1 QIPRIPDAL DVVPQCILTS ASPTPLFPSS SPANGDACI FGTIVRSGE
 51 SFQSCSKYOC TGLDAGVGM PLASMDVRLP SPDCFPFRV KLPKSCCEEM
 101 VCDSPKQGV IGPASVRSR FLXRVVILS QGSPNCADR TGEIPYPGVD
 151 HGVCVLCRS LPTGRHWPR PNYDXQLPG PTEWSACSK TCGNGYSTRV

201 TNDNA

11AA SEQUENCE 1.0
 ID R30618 standard; Protein: 106 AA.
 AC R30618-1993 (first entry)
 DT 19-MAY-1993
 DE Polypeptide coded by Korean HCV cDNA fragment NS2-LB3.
 KM KHCY; diagnosis; vaccine; NS2 region.
 OS Korean hepatitis C virus.
 PN EP-92-118-A.
 PD 07-JAN-1993.
 PF 10-JUN-1992; 109753.
 PR 06-AUG-1991; KR-008510.
 PA (LDCR) LOCKY LTD.
 PI Cho JM, Choi DY, Kim CH, Kim ST, Lee YB, Lim KJ, Park YW,
 PI So HS, Yang JY;
 DR WP1: 93-001883/01.
 DR N-PSDB: Q33284.
 PT DNA and polypeptide(s) from a new type of hepatitis C virus (KHCY)
 PT - for diagnosing and vaccinating against KHCY infections
 PS Disclosure; Fig 8; 11pp; English.
 CC The polypeptide is that encoded by cDNA fragment NS2-LB3 obtd. from
 CC Korean hepatitis C virus (KHCY) RNA extracted from the sera of
 CC hepatitis C patients. It or its fragments may be used in a specific
 CC and accurate method for detecting KHCY antibodies in the serum of
 CC hepatitis C patients. Antibodies directed against the polypeptide
 CC are useful for the purification of KHCY antigens and for the
 CC development of an improved diagnostic to detect KHCY antigens in a
 CC sample. The polypeptide may also be used in a vaccine for treatment
 CC and prevention of KHCY infection at a dosage of 5-200 ug/peptide.
 SO Sequence 106 AA.

R30618 Length: 106 February 11, 2000 15:48 Type: P Check: 7998

1 LILSPYKY FLARIMWLQ YFTRRAHD QWVPIPLNVR GGDALILIA
 51 CAVHPEIFD ITRKLLAIFG PLWVLAQGIT RVPYFWRAG LIRACMLARK
 101 VAGGHY

11AA SEQUENCE 1.0
 ID R32192 standard; Protein: 246 AA.
 AC R32192-1993 (first entry)
 DT Sequence encoded by the acetoacetyl-CoA reductase (phbB) gene of the
 DE polyhydroxybutyrate (PHB) operon
 KM Operon: polyhydroxyalkanoate: acetoacetyl-CoA reductase.
 OS Alcaligenes eutrophus.
 PN WO9302187-A.
 PD 04-FEB-1993.
 PF 13-JUL-1992; U05786.
 PR 19-JUL-1991; US-732243.
 PA (UYMA-) UNIV MADISON JAMES.
 PA (UNMS) UNIV MICHIGAN STATE.
 PI Dennis DE, Poirier Y, Somerville CR;
 DR WP1: 93-058785/07.
 DR N-PSDB: Q36660.
 PT Transgenic plants producing poly(hydroxy-alkanoate polymer(s)) -
 PT obtd. by transformation with DNA encoding 3-ketothiolase,
 PT acetoacetyl-CoA reductase and PHA synthase
 PS Disclosure; Fig 2; 70pp; English.
 CC The nucleotide sequence of the PHB operon was obtained from James, B.
 CC Holler, J. and Dennis, D. in Daves, E.A. (ed.) Novel Biodegradable
 CC Polymers, Kluwer Academic Publishers, 175-190 (1990). It contains
 CC the genes from PHB synthase, 3-ketothiolase and acetoacetyl-CoA
 CC reductase. The inventors claim a transgenic plant material congt.
 CC foreign DNA encoding a peptide which exhibits 3-ketothiolase activity,
 CC pref. where the DNA is an open reading from between nucleotides
 CC 3696-3877 (phb A gene), 842-2611 (phb C gene) or 3952-4692 (phb B
 CC gene) of the Alcaligenes eutrophus PHB operon.
 SO Sequence 246 AA.

R32192 Length: 246 February 11, 2000 15:48 Type: P Check: 2829 ..

1 MTORLAVVTG GNGIGTALC ORLAKDGRV VAGCGSPR REKMLEOQKA

51 LGPEFIASEG NVADMSTKT AFDKXSEVG EVDVLNMG IRDVFYFKM

101 TRAMDAVID TMTSLFNT KQVIGMDR GMGRIVNMS VAGCGGCGGQ

151 TMTSTAKAGL HGFTMLAOE VAIKGVYNT VSPGYATDM VAIKQDYLD

201 KIVATIPYKR LGPEFIASI CAMLSSESG FSTGADPSLN GGLHNG

11AA-SEQUENCE 1.0

ID R37294 standard; protein; 248 AA.

AC R37294;

DE 13-SEP-1993 (first entry)

DE Plant type I RIP luffin.

KW Type I ribosome-inactivating protein; ricin; gelonin; momordin; immunocytogate; autoimmune disease; cell killing; toxin; locofah.

OS Luffa

PN WO9309130-A.

PD 13-MAY-1993.

PF 04-NOV-1992; U09487.

PR 04-NOV-1991; US-787567.

PR 19-JUN-1992; US-901707.

PA (XOMA) XOMA CORP.

PI Bernhard SL, Better MD, Carroll SF, Lane JA, Lei SP;

DR WPI: 93-167617/20.

PT Analogues of type I ribosome inactivating protein - useful as cytotoxic agents, immunotoxins for treating autoimmune diseases, cancer, graft versus host disease and selective cell killing in vivo

PS Disclosure: Page 95-96; 163pp. English.

CC The invention covers analogues of type I RIPS. Luffin is a type I RIP and the analogues of the invention have a cysteine available for intermolecular disulphide bonding at an amino acid position corresp. to a position not naturally available for bonding; the cys residue is located in the C-terminal region of the analogue between a position corresp. to amino acid 251 and the C-terminus of ricin A chain. The analogues are pref. joined via a disulphide linkage to a molecule which specifically binds to a target cell, e.g. an antibody fragment.

CC Sequence 248 AA;

R37294 Length: 248 February 11, 2000 15:48 Type: P Check: 7673 ..

1 DVFSLSGSS STSYSKFID LKALPNSGT VYNLTLLSS ASGASRYTLM

51 TLSNYDKAI TVADVSOYL IMGYLVNST YFNEPDAKL ASQYVFKGST

101 IVTLPYSGNY EKLQTAAGRI REKIPLGPPA LDSALTTIFH YDSTAAAF

151 LVILQTAEA SREKYIEGDI IERISKQVP SLATISLENS LMSALSKQIO

201 LAQNNNGTK TPVYITDDKG QRELETNTS KVTTKNIQLL LMYKQNV

11AA-SEQUENCE 1.0

ID R39259 standard; protein; 391 AA.

AC R39259;

DE 18-NOV-1993 (first entry)

DE Human somatostatin receptor-1.

KW Human somatostatin receptor; SSFR-1; SSFR-2; SSFR-3; tumour; pancreas; islet; promoter; transformation; host cell.

OS Homo sapiens.

PN WO9313130-A.

PD 08-JUL-1993.

PF 30-DEC-1992; U11327.

PR 31-DEC-1991; US-816283.

PA (ARCH-) ARCH DEV CORP.

PI Bell GI, Selino S, Yamada Y;

DR WPI: 93-227272/28.

DR N-PSDB; Q45653.

PT Somatostatin receptors useful for diagnosis of tumours - also

PT useful for screening candidate somatostatin receptor agonists and antagonists

PS Claim 3; Page 63-64; 94pp; English.

CC The sequences given in R39259, R39261 and R39263 represent the human somatostatin receptors (SSFR)-1, SSFR-2 and SSFR-3. The DNA encoding these proteins was isolated from total human pancreatic islet DNA.

CC These DNA sequences may be placed under the control of a suitable promoter and used to transform a host cell. The DNA sequences and these proteins may be used in screening assays for testing candidates including agonists and antagonists of SSFR polypeptides. The assays may be used to discriminate candidate substances with desirable properties specific to SSFR polypeptides. The isolated substances may be used in a wide range of applications eg. diagnosis of various human tumours. Fragments of these DNA sequences may be used as probes in the isolation of other SSFR-encoding clones.

CC Sequence 391 AA;

R39259 Length: 391 February 11, 2000 15:48 Type: P Check: 8652 ..

1 MFPNGTASSP SSSPSPSPSG CGEGGSRGP GAGADGME PGRNAQNGT

51 LSEGGSSAIL ISFIYSVCL VGICGNSWVI YVILRYANM TRINITYLNL

101 AIADFLMKS VPFLVSTLL RHMPFGALIC RLIVSDAVN MFTSIVCLIV

151 LSVDRYVAVV HPIKARYRR PIVAKVYVNG VWVLSLVYL PIVFESRTAA

201 NSDGTACNM IMPEPAQRL VGEVLYTFIL GFLPVGAIK ICYVLIHAKM

251 RMVALAGNM QRRSERKIT LWMVMVWVF VICMMPFVY QLVNVAEED

301 DAVSGLSVI LGYNSCANP ILVGFSLDNF KRSPORLCL SWMDNAEEP

351 VDYATALKS RAYVEDPQ ENLESGVFR NGTCISRTT L

11AA-SEQUENCE 1.0

ID R39260 standard; protein; 391 AA.

AC R39260;

DE 18-NOV-1993 (first entry)

DE Murine somatostatin receptor-1.

KW Mouse; somatostatin receptor; SSFR-1; SSFR-2; SSFR-3; tumour; pancreas; islet; promoter; transformation; host cell.

OS Mus musculus.

PN WO9313130-A.

PD 08-JUL-1993.

PF 30-DEC-1992; U11327.

PR 31-DEC-1991; US-816283.

PA (ARCH-) ARCH DEV CORP.

PI Bell GI, Selino S, Yamada Y;

DR WPI: 93-227272/28.

DR N-PSDB; Q45654.

PT Somatostatin receptors useful for diagnosis of tumours - also useful for screening candidate somatostatin receptor agonists and antagonists

PS Claim 3; Page 65-66; 94pp; English.

CC The sequences given in R39260, R39262 and R39264 represent the murine somatostatin receptors (SSFR)-1, SSFR-2 and SSFR-3. The DNA encoding these proteins was isolated from total murine pancreatic islet DNA.

CC These DNA sequences may be placed under the control of a suitable promoter and used to transform a host cell. The DNA sequences and these proteins may be used in screening assays for testing candidates including agonists and antagonists of SSFR polypeptides. The assays may be used to discriminate candidate substances with desirable properties specific to SSFR polypeptides. The isolated substances may be used in a wide range of applications eg. diagnosis of various human tumours. Fragments of these DNA sequences may be used as probes in the isolation of other SSFR-encoding clones.

CC Sequence 391 AA;

R39260 Length: 391 February 11, 2000 15:48 Type: P Check: 8110 ..

1 MFPNGTASSP SSSPSPSPSG CGEGGSRGP GAGADGME PGRNAQNGT


```

51 LSEGGSSAIL ISFIYSVCL VGLGNSKVI YVILRYAKK TATNITLNL
101 AIADELLMS VPELVSTLL RHPFGALLC RLIVSDAVN MFSITCLTV
151 LSVDRYAVV HPKARVRR PIVAKVNLG VWLISLVIL PIVFSRTAA
201 NSDGVACNM LMPEPARML VGFVLYTFLM GELLPGALC LCVYLIAMK
251 RVVALKAGM ORKRSERKIT LMYMVAVVF VICMPEYVY QLVNFAEOD
301 DATVSQLSVT LGVANSQAN IIXGLSDNF KRSQRIICL SMDNAASEP
351 VDYVATALKS RAYVEDFOP ENLESGVFR NGTCASRIST L

!!AA SEQUENCE 1.0
ID R42317 standard; Protein; 345 AA.
AC R42317;
DE 22-MAY-1994 (first entry)
KW EBV VCA-p40 from ORF Bdrf1.
KW Viral capsid antigen; Epstein Barr Virus; viral genome; antibodies;
KM diagnosis.
OS Epstein-Barr virus.
PN AU9335152-A.
PD 16-SEP-1993.
PE 12-MAR-1993; 035152.
PR 13-MAR-1992; BP-200721.
PA (ALKU) AKZO NV.
PI Middelorp JM.
DR WPI: 93-345368/44.
DR N-PSDB: 051019.
PT Peptide(s) and nucleic acid sequences related to EBV - for
PT detecting EBV in samples
PS Claim 3: Page 40; 62pp; English.
CC The sequence is that of Epstein Barr Virus viral capsid antigen
CC VCA-p40 encoded by the EBV open reading frame Bdrf1. This antigen is
CC immunochemically reactive with antibodies to EBV and can be used for
CC the diagnosis of EBV in a sample.
CC See also R42316-35.
SQ Sequence 345 AA;

R42317 Length: 345 February 11, 2000 15:48 Type: P Check: 3489

1 MLSGNAGEGA TACGSAAG ODLISVPNT FMTLOTIND NKPRPOTLP
51 YAAPLPESH QAITAPSYG PGAGAVAPAG GYFTSPGGY AGPAGDPGA
101 FLAMDATHV PHRPPEAYE GLPGFPP PPVPYGSML RADYVAPSR
151 SKRRRDPEE DEBGGLFPG EDATIXRDI AGLSKVNEL QHTLALRRE
201 TISYHTGVG YCPQGFCTY HSGPYGFQPH QSYEVPRYV HPPPPTSHQ
251 AAQAOPPPPG TQAPFAHVA ESTIPBAGAA RNSGPRDIN PQOPTTEGHH
301 RGRKLVOASA SGVAOSKEPT TRKANSVAH LKSTICEELL MKRYA

!!AA SEQUENCE 1.0
ID R52667 standard; Protein; 303 AA.
AC R52667;
DE 09-AUG-1994 (first entry)
KW Equine herpesvirus US2 gene.
KW Equine herpesvirus; US2; vaccine; antigen; protection; prophylaxis;
KW prevention.
OS Equine herpesvirus (Strain: Dutta; individual isolate: S-1EHV-000).
PN W09403628-A.
PD 17-FEB-1994.
PE 06-AUG-1993; 007424.
PR 07-AUG-1992; US-926784.
PA (CYTR) SYNPRO CORP.
PI Cochran MD.
DR WPI: 94-064715/08.
DR N-PSDB: Q56616.
PT New recombinant equine herpes viruses - used to prepare vaccines

```

```

PT to protect horses from infectious equine herpes virus
PS Example 1: Page 83-84; 153p; English.
CC Recombinant equine herpesvirus (EHV) have a foreign DNA sequence
CC inserted into their genomes. The inserted sequence is a piece of
CC foreign, double stranded DNA which encodes an RNA molecule which
CC does not naturally occur in the animal into which the EHV is
CC introduced. The foreign DNA sequence is introduced into the EHV via
CC a homologous vector. The recombinant EHV's are used to prepare
CC vaccines to protect horses from infectious EHV. This is the US2
CC gene product from a recombinant EHV. Recombinant attenuated EHV's
CC can be produced by removing the US2 gene
SQ Sequence 303 AA;

R52667 Length: 303 February 11, 2000 15:48 Type: P Check: 5723

1 MGVLIITVY VDRHKALEN SSIDVGHLM EFLSRQCVL ASEPLGPIV
51 VRSADLYRFS SLTLTPKAC RPIVTRGAT AILDENGVV YHEDRMGYSI
101 EMLSVLSGIN HLNSSLINQ PPHLMVLGAA DLCKPFDLI PGKRMVYAE
151 IADFEHKSQ PPVYCGKLE TTPWTVEHN HPLKLAAGS EDTVVGECGF
201 SKHSNSIVR PPTVRYVIA VDPARLREI PARGPLPR RSESGMRAP
251 RRRSRAPAA RSTAAATPP RGDPRAPAA RRAQDTWME RLNGVFGRT
301 STR

!!AA SEQUENCE 1.0
ID R49296 standard; Protein; 25 AA.
AC R49296;
DE 16-SEP-1994 (first entry)
KW Invariant chain (Ii) position 97-121.
KW Naturally-occurring; immunomodulatory protein; human; therapy; class I;
KW major histocompatibility complex; class II; allotype; type I diabetes;
KW autoimmune disease; rheumatoid arthritis; T-cell-mediated response;
KW multiple sclerosis; transplant rejection; vaccine; MHC.
OS Homo sapiens.
PN W09404171-A.
PD 03-MAR-1994.
PR 11-AUG-1993; 007545.
PR 11-AUG-1992; US-925460.
PR 15-JUN-1993; US-925460.
PA (HARD) HARVARD COLLEGE.
PI Chaz RM, Hedley ML, Stern LJ, Strominger JL, Urban RG;
PI Vignali DA.
DR WPI: 94-082825/10.
PT Novel immunomodulatory peptide(s) and nucleic acids - useful for
PT treatment of autoimmune diseases, transplant rejection and for
PT vaccination; Page 36; 139pp; English.
PS The sequences given in R49291-505 and R49981-7038 represent peptide
CC fragments of naturally occurring immunomodulatory proteins. These
CC fragments are between 10-30 residues in length and bind to a human
CC major histocompatibility complex (MHC) class II allotype. These
CC peptides may be used for therapy of autoimmune diseases, such as
CC type I diabetes, rheumatoid arthritis and multiple sclerosis, and to
CC reduce transplant rejection. They may also be used for vaccination
CC providing an exclusively T-cell mediated response which can be
CC class I or class-II based, or both, depending on the length and
CC character of the immunogenic peptides.
SQ Sequence 25 AA;

R49296 Length: 25 February 11, 2000 15:48 Type: P Check: 4901

1 LPPRPVSK MRNATPLMO ALPMG

!!AA SEQUENCE 1.0
ID R49297 standard; Protein; 24 AA.
AC R49297;
DE 16-SEP-1994 (first entry)
KW Invariant chain (Ii) position 97-120.

```

KW Naturally-occurring; immunomodulatory protein; human; therapy; class I;
 KW major histocompatibility complex; class II; allotype; type I diabetes;
 KW autoimmune disease; rheumatoid arthritis; T-cell-mediated response;
 KW multiple sclerosis; transplant rejection; vaccine; MHC.
 OS Homo sapiens.
 PN WO9404171-A.
 PD 03-MAR-1994.
 PF 11-AUG-1993; U07545.
 PR 11-AUG-1992; US-925460.
 PR 15-JUN-1993; US-925460.
 PA (HARD) HARVARD COLLEGE.
 PI Chicz RM, Hedley ML, Stern LJ, Strominger JL, Urban RG;
 PI Vignali DA;
 DR WPI: 94-082825/10.
 PT Novel immunomodulatory peptide(s) and nucleic acids - useful for
 PT treatment of autoimmune diseases; transplant rejection and for
 PT vaccination.
 PS Disclosure: Page 36; 139pp; English.
 CC The sequences given in R49291-505 and R46981-7038 represent peptide
 CC fragments of naturally-occurring immunomodulatory proteins. These
 CC fragments are between 10-30 residues in length and bind to a human
 CC major histocompatibility complex (MHC) class II allotype. These
 CC peptides may be used for therapy of autoimmune diseases, such as
 CC type I diabetes, rheumatoid arthritis and multiple sclerosis, and to
 CC reduce transplant rejection. They may also be used for vaccination
 CC providing an exclusively T-cell-mediated response, which can be
 CC class I or class-II based, or both, depending on the length and
 CC character of the immunogenic peptides.
 SQ Sequence 24 AA;

R49297 Length: 24 February 11, 2000 15:48 Type: P Check: 3126 ..

1 LPPKPVSK MRMATPLMQ ALPM

IIAA-SEQUENCE 1.0
 ID R49299 standard; Protein: 23 AA.
 AC R49299;
 DT 16-SEP-1994 (first entry)
 DE Invariant chain (I1) position 97-119.
 KW Naturally-occurring; immunomodulatory protein; human; therapy; class I;
 KW major histocompatibility complex; class II; allotype; type I diabetes;
 KW autoimmune disease; rheumatoid arthritis; T-cell-mediated response;
 KW multiple sclerosis; transplant rejection; vaccine; MHC.
 OS Homo sapiens.
 PN WO9404171-A.
 PD 03-MAR-1994.
 PF 11-AUG-1993; U07545.
 PR 11-AUG-1992; US-925460.
 PR 15-JUN-1993; US-925460.
 PA (HARD) HARVARD COLLEGE.
 PI Chicz RM, Hedley ML, Stern LJ, Strominger JL, Urban RG;
 PI Vignali DA;
 DR WPI: 94-082825/10.
 PT Novel immunomodulatory peptide(s) and nucleic acids - useful for
 PT treatment of autoimmune diseases; transplant rejection and for
 PT vaccination.
 PS Disclosure: Page 36; 139pp; English.
 CC The sequences given in R49291-505 and R46981-7038 represent peptide
 CC fragments of naturally-occurring immunomodulatory proteins. These
 CC fragments are between 10-30 residues in length and bind to a human
 CC major histocompatibility complex (MHC) class II allotype. These
 CC peptides may be used for therapy of autoimmune diseases, such as
 CC type I diabetes, rheumatoid arthritis and multiple sclerosis, and to
 CC reduce transplant rejection. They may also be used for vaccination
 CC providing an exclusively T-cell-mediated response, which can be
 CC class I or class-II based, or both, depending on the length and
 CC character of the immunogenic peptides.
 SQ Sequence 23 AA;

R49299 Length: 23 February 11, 2000 15:48 Type: P Check: 1278 ..

1 LPPKPVSK MRMATPLMQ ALP

IIAA-SEQUENCE 1.0
 ID R49438 standard; Protein: 49 AA.
 AC R49438;
 DT 16-SEP-1994 (first entry)
 DE M19 gene HLA-DRA alpha chain/I1 24 residue peptide fusion.
 KW Naturally-occurring; immunomodulatory protein; human; therapy; class I;
 KW major histocompatibility complex; class II; allotype; type I diabetes;
 KW autoimmune disease; rheumatoid arthritis; T-cell-mediated response;
 KW multiple sclerosis; transplant rejection; vaccine; MHC.
 OS Homo sapiens.
 PN WO9404171-A.
 PD 03-MAR-1994.
 PF 11-AUG-1993; U07545.
 PR 11-AUG-1992; US-925460.
 PR 15-JUN-1993; US-925460.
 PA (HARD) HARVARD COLLEGE.
 PI Chicz RM, Hedley ML, Stern LJ, Strominger JL, Urban RG;
 PI Vignali DA;
 DR WPI: 94-082825/10.
 PT Novel immunomodulatory peptide(s) and nucleic acids - useful for
 PT treatment of autoimmune diseases; transplant rejection and for
 PT vaccination.
 PS Disclosure: Page 93; 139pp; English.
 CC The sequences given in R49291-505 and R46981-7038 represent peptide
 CC fragments of naturally-occurring immunomodulatory proteins. These
 CC fragments are between 10-30 residues in length and bind to a human
 CC major histocompatibility complex (MHC) class II allotype. These
 CC peptides may be used for therapy of autoimmune diseases, such as
 CC type I diabetes, rheumatoid arthritis and multiple sclerosis, and to
 CC reduce transplant rejection. They may also be used for vaccination
 CC providing an exclusively T-cell-mediated response, which can be
 CC class I or class-II based, or both, depending on the length and
 CC character of the immunogenic peptides.
 SQ Sequence 49 AA;

R49438 Length: 49 February 11, 2000 15:48 Type: P Check: 4300 ..

1 MAISGVPVIG FFIIVALMSA QESWALPKRP KPVSKMMAT PLMQALPM

IIAA-SEQUENCE 1.0
 ID R49589 standard; Protein: 49 AA.
 AC R49589;
 DT 15-SEP-1994 (first entry)
 DE Sequence of HLA-DR alpha chain leader peptide linked to the
 DE amino terminus of a blocking peptide fragment of human invariant
 DE chain I1.
 KW M19 gene; HLA-DR alpha leader peptide; chain I1.
 OS Synthetic.
 FH Key
 FT peptide
 FT Location/Qualifiers
 FT 1..25
 FT /label= HLA-DR alpha leader peptide
 FT 26..49
 FT /label= chain I1
 PN WO9404557-A.
 PD 03-MAR-1994.
 PF 11-AUG-1992; U06692.
 PR 11-AUG-1992; WO-006692.
 PA (HARD) HARVARD COLLEGE.
 PI Chicz RM, Hedley ML, Stern LJ, Strominger JL, Urban RG;
 PI Vignali DA;
 DR WPI: 94-083102/10.
 PT New peptide binding to MHC class II allotype - useful for
 PT treating autoimmune diseases; transplant rejection and for
 PT immunisation.
 PS Example: Figure 3B; 60pp; English.
 CC The inventors claim a nucleic acid encoding a polypeptide
 CC comprising an AA sequence identical to that of a segment of a
 CC naturally-occurring human protein which binds to a human
 CC major histocompatibility complex (MHC) class II allotype and
 CC a second sequence substantially identical to human I1. In order
 CC to construct such minigenes, overlapping synthetic oligos
 CC were used to generate the leader peptide/blocking peptide

CC minigenes (see Q44333/R49588, Q44334/R49589). The resulting
CC constructs were cloned into pGEM-2.
SQ Sequence 49 AA;

R49589 Length: 49 February 11, 2000 15:48 Type: P Check: 4300 ..

1 MAISCVPLIG FFIIVLMSA QESWALPKRP KPVSKMMAT PLIMOLPM

!!AA_SEQUENCE 1.0
ID R50081 standard; Protein: 352 AA.

DE 24-NOV-1994 (first entry)
DT NANBH virus antigenic fragment #13.
CC Antigen; structural; non-structural; non A non B hepatitis virus;
KW NANBH; NANBH; patient; plasma; diagnosis; detection; carrier; ss.
OS Non A non B hepatitis virus.
PN J06070778.A.
PD 15-MAR-1994.
PF 01-JUN-1993; 156087.
PR 10-JUL-1992; JP-207391.
PA (KOKU-) KOKUSAI SHIYAKU KK.
PA (SANKU-) SANMA KAGAKU KENKYUSHO CO.
PA (TOFU-) TONEN CORP.
PA (TOKR-) 2H TOKYO RINSHO IGAKU SOGO KENKYUSHO.
DR N-PSDB; Q58826.
PT Nucleic acid fragment coding non-A non-B hepatitis virus antigen
PT - useful in diagnosis of NANB patient and detection of virus
PT carrier
PS Claim 26; Page 31-32; 37pp; Japanese.
CC The sequences given in R50068 and R50070-82 represent antigens of
CC structural and non-structural regions of non A non B hepatitis virus
CC (NANBH). The cDNA encoding these sequences were derived from the
CC plasma of a NANBH patient by recombinant DNA techniques. These
CC fragments are useful for the diagnosis of NANBH patients and the
CC detection of NANBH carriers.
SQ Sequence 352 AA;

R50081 Length: 352 February 11, 2000 15:48 Type: P Check: 3747 ..

1 LBRHNKYYC TTSSASJRA KKVTFPMQV LDAHVSYLK DIKLSKVS

51 AALLTLEEC RLPFPHSARS KYGFGAKVR SLSGAVNHI KSVMKLLED
101 PQTPIPTITM AKNEVFCVDP TRGGRKPARL IYVPLGVAV CEMALYDVT
151 QRLPQAVMGA SYGFQYSPAQ RVEFLKAMA DKQDMGFSY DTRCFDSITY
201 ERDIRTEESI YQACSLPEEA RTAIRSLTER LYVGPMPNS KGQACGYRRC
251 RASGVITSTM GNTTCYVKA LAACKRAGIV APTMLVCGDD LVYSESQGT
301 EEDERNLRAF TEAMTRYSAP PGDPPEPYD LELITGSSN VSYAISPQGR
351 RR

!!AA_SEQUENCE 1.0
ID R50082 standard; Protein: 215 AA.

DE 24-NOV-1994 (first entry)
DT NANBH virus antigenic fragment #14.
CC Antigen; structural; non-structural; non A non B hepatitis virus;
KW NANBH; NANBH; patient; plasma; diagnosis; detection; carrier; ss.
OS Non A non B hepatitis virus.
PN J06070778.A.
PD 15-MAR-1994.
PF 01-JUN-1993; 156087.
PR 10-JUL-1992; JP-207391.
PA (KOKU-) KOKUSAI SHIYAKU KK.
PA (SANKU-) SANMA KAGAKU KENKYUSHO CO.
PA (TOFU-) TONEN CORP.
PA (TOKR-) 2H TOKYO RINSHO IGAKU SOGO KENKYUSHO.
DR WPI; 94-128677/16.

DR N-PSDB; Q58827.
PT Nucleic acid fragment coding non-A non-B hepatitis virus antigen
PT - useful in diagnosis of NANB patient and detection of virus
PT carrier

PS Claim 28; Page 32-33; 37pp; Japanese.
CC The sequences given in R50068 and R50070-82 represent antigens of
CC structural and non-structural regions of non A non B hepatitis virus
CC (NANBH). The cDNA encoding these sequences were derived from the
CC plasma of a NANBH patient by recombinant DNA techniques. These
CC fragments are useful for the diagnosis of NANBH patients and the
CC detection of NANBH carriers.
SQ Sequence 215 AA;

R50082 Length: 215 February 11, 2000 15:48 Type: P Check: 108 ..

1 RYSAPQDDP REYDELEIT SCSSNSVAL SPQGRRRYL SRDPTPIAR

51 AMETVARNHP VNSWGNITQ YAPTIWVWV LMTHEFPVLI AODLDDNLN
101 FENWGSYISV SPLDPAIEE RLHGDLAFSL HRYTHHELTR VASALRLGA
151 PPLRAKMSRA RAVRASLISO GGRAAVGGRY LNNMAVKIKL KLPLPPEARL
201 LDLSMFTVG AGGCD

!!AA_SEQUENCE 1.0
ID R53251 standard; Protein: 56 AA.

DE 09-DEC-1994 (first entry)
DT Consensus sequence of signal peptide of small subunit of RUBISCO.
KW Glycogen synthase; starch; gene expression; recombinant;
KW transgenic plant; glucose-1-phosphate adenylyl transferase;
KW biosynthesis; metabolism; metabolic pathway.
OS Synthetic.
PN W09409144.A.
PD 28-APR-1994.
PF 14-OCT-1992; G01881.
PR 14-OCT-1992; AU-026964.
PR 14-OCT-1992; WO-G01881.
PA (GENE) ZENEGA LTD.
PI Bird CR, Fentem PA, Keeling PL, Singletary G;
DR WPI; 94-151328/18.
PT New plants with altered starch synthesis ability - obld. by
PT incorporating a donor gene specifying an enzyme involved in
PT starch or glycogen biosynthesis
PS Disclosure; Page 31; 91pp; English.
CC The glycogen synthase gene may be incorporated into a plant genome
CC to produce a plant with altered starch synthesizing ability. Other
CC genes encoding enzymes involved in the starch or glycogen
CC biosynthetic pathways may also/optionally be used. The new plants
CC have an improved capacity to produce starch at elevated or lowered
CC temperatures and/or an ability to synthesise starch with an altered
CC fine structure. They can produce higher starch yields at certain
CC temperatures and/or produce starch with an improved quality. The
CC glycogen synthase gene construct used requires the presence of an
CC amyloplast transit peptide to ensure its correct localisation in the
CC amyloplast. This is the consensus sequence of the transit peptide of
CC the small subunit of RUBISCO from many genotypes.
SQ Sequence 56 AA;

R53251 Length: 56 February 11, 2000 15:48 Type: P Check: 2035 ..

1 MASSNLSAA VATRINPAQA SMVAPFTGLK SAAFPVSRKO NLDITISAN

51 GGRVQC

!!AA_SEQUENCE 1.0
ID R53646 standard; Protein: 380 AA.

AC R53646;
DE 18-JAN-1995 (first entry)
DT c-fos gene product.
KW c-jun; c-fos; cancer; tumour; metastasis; treatment; prophylaxis;
KW vaccine; prevention; lymphocytes; inhibition; antigen; antitumour.

OS Homo sapiens.
 PN EP-59077-A.
 PD 01-JUN-1994.
 PE 28-OCT-1993: 117519.
 PR 29-OCT-1992: US-968415.
 PA (YEDA) YEDA RES & DEV CO LTD.
 PI Eisenbach L, Feldmann M;
 DR WPI: 94-169430/21.
 DR N-PSDB: 063815.
 PT New antitumour vaccines partic. for cancer treatment - comprises
 PT tumour cells having inserted c-fos and opt. c-jun genes, or
 PT antigens expressed by these genes
 PS Disclosure: Page 54-55; 68pp; English.
 CC The c-fos gene can be transfected alone or co-transfected with the
 CC c-jun gene (See 063813) into tumour cells to provide a novel
 CC anti-tumour cellular vaccine. The anti-tumour vaccine comprises
 CC antigens expressed by the c-fos and c-jun genes. The vaccines
 CC stimulate cytotoxic T-lymphocytes to produce an anti-tumorigenic and
 CC anti-metastatic immune response. They can be used for controlling
 CC tumorigenicity and metastatic properties of tumours. They are used
 CC particularly for treating a patient suffering from cancer to prevent
 CC and/or inhibit the development of metastasis.
 SQ Sequence 380 AA;

R53646 Length: 380 February 11, 2000 15:48 Type: P Check: 4677 ..

1 MMSGFNAY EASSRCSA SPAGDSLSY HSPADFSM GSPVNAQDFC
 51 TDLAVSSANF IPTVTAISTS PDLQWLVOA LVSSVAPSQT RAPHFGVPA
 101 PSAGATSRAG VKTWTGGR QSIGRGKVE QLSPEEERK RIRERKMA
 151 AAKCRNRRE LDTLQAFED QLEDEKSAQ TEIANLKER EKLEFILA
 201 RPACIPDDL GFPEMSVA LDTLGLPEV ATPSEEAFT LPLNDEPK
 251 PSVEPKSIS SMELKTEPD DFLFPASSR SGESEARAV DMDLGSFYA
 301 ADMEPLHSGS LGMGPMA TEL EPLCTPVVC TPSCATYSS FVTFPEADS
 351 FPSCAAHKR GSSSNPSSD SLSPITLAL

!!AA_SEQUENCE 1.0

ID R53748 standard; Protein; 355 AA.

AC R53748;

DT 06-FEB-1995 (first entry)

DE Seven transmembrane receptor (V28).

KW Primer: seven transmembrane receptor; receptor; amplification; PCR;

KW polymerase chain reaction.

OS Synthetic.

PH Key Location/Qualifiers

FT domain 26..56

FT domain /label= Transmembrane domain.

FT domain /label= 68..92

FT domain /label= Transmembrane domain.

FT domain 107..125

FT domain /label= Transmembrane domain.

FT domain 146..167

FT domain /label= Transmembrane domain.

FT domain 197..219

FT domain /label= Transmembrane domain.

FT domain 232..253

FT domain /label= Transmembrane domain.

FT domain 273..297

FT domain /label= Transmembrane domain.

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

PT DNA encoding seven transmembrane receptors - used to develop
 PT prods. for use as therapeutic or diagnostic agents for conditions
 PT involving the receptors.
 PS Example 5; Page 65-66; 100pp; English.
 CC Two synthetic oligonucleotides (066165, 066166) were annealed and
 CC radiolabelled to produce a V28 specific probe. The probe was
 CC denatured by boiling for two minutes and then hybridised to the
 CC human placenta genomic library. Several strong hybridising signals
 CC were identified and plaque purified. A full length genomic clone
 CC of a full length genomic clone encoding this seven transmembrane
 CC receptor was deduced.
 SQ Sequence 355 AA;

R53748 Length: 355 February 11, 2000 15:48 Type: P Check: 2863 ..

1 MDQFESVTE NFEYDILA CYIGDIYFG TVFLSIFYSV IFAIGLVNL
 51 LVVFALINSK KRSVTDIYL LNLASDLF VATLFPWTHY LINKGLHNA
 101 MCFETAFEF IGFGSIFPI TVSIDRYLA IYLA NSMN RYQGVYIS
 151 LGVMAAIIY AAPQFMFTK KENDCLGDP EYLQETWPL RNVETNLF
 201 LPLLLMSYC YFRIGTLFS CKNKKRKA KILLVIVF FLEWTPYVM
 251 IFLETLKLYD FFPQDMRKD LRLALVTE VAFSHCCINP LIYAFAGEK
 301 RRYLYHLCK CLAVLCGRSV HYDFSSESO RSRHGSVLSS NITYHTSDGD
 351 ALLLL

!!AA_SEQUENCE 1.0

ID R58556 standard; Protein; 175 AA.

AC R58556;

DT 28-FEB-1995 (first entry)

DE Heparin-binding secretory transforming factor 2.

KW Heparin-binding growth factor family; fibroblast growth factor;

KW FGF; hst-2; heparin-binding secretory transforming factor 2;

KW drugs; reagents; growth promoting activity; increase platelet number;

KW angiogenesis promoting activity; treatment; burns; ulcers.

OS Synthetic.

PH Key Location/Qualifiers

FT misc_difference 1

FT /note= "may be absent"

FT misc_difference 2..30

FT /note= "N terminal fragment as given in the

FT specification"

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

R58556 Length: 175 February 11, 2000 15:48 Type: P Check: 4528 ..

1 MLAGIAGVN WSGYLVGK RQRRLCNGV IGFIHQVLEPD GRISGTHEN

51 PYSLEISTV ERGVSLFV RSLFVAMNS KGLYATPSF QBECKFRET
 101 LPNNYAES DLYOGTYIAL SKYGRVKRGS KVPSPMTVTH FLPRIVSPS
 151 AGTRANNTIL DSRGWGTLIS RSRAG

11AA_SEQUENCE 1.0
 ID R63592 standard; Peptide; 79 AA.
 AC R63592:
 DT 09-MAY-1995 (first entry)
 DE MAP-kinase-phosphatase Ctl100 partial sequence.
 KW MAP-kinase-inhibitor; MAPK; MAP-kinase-phosphatase; Ctl100;
 RW neurological disease; proliferative disorder; inflammation;
 OS cardiovascular disease.
 PS Homo sapiens.
 PN MO9423039-A.
 PD 13-OCT-1994.
 PE 31-MAR-1994; G00694.
 PF 07-MAR-1993; GB-002750.
 PR 10-FEB-1994; GB-002573.
 PA (CANC-) CANCER RES INST.
 PI (CANC-) INST CANCER RES ROYAL CANCER HOSPITAL.
 DR Ashworth A, Hughes DA, Marshall CJ;
 PI WPI: 94-333199/41.
 PT Screening for a substance which is an inhibitor of mammalian MAP
 kinase - and is used in the mfr. of a medicament for the
 PT treatment of e.g. an inflammatory disorder or neurological
 PT disease.
 PS Disclosure: Fig. 7a: 53pp; English.
 CC STY2, STY3, STY4 and STY5 cDNAs (given in 072858-61), encoding
 CC peptides having MAP-kinase-phosphatase activity, were derived from
 CC human squamous A431 cell polyA+ mRNAs. Residues 177-255 of the
 CC encoded STY2-5 peptides were deduced (R63593-96) and demonstrated
 CC to show homology with the corresponding region of the previously
 CC isolated human Ctl100 gene product (R63592).
 SQ Sequence 79 AA.

R63592 Length: 79 February 11, 2000 15:48 Type: P Check: 2807

1 ILPELYIGSA YHASRKMDL ALGITALINV SANCNHEEG HYQKSIPEV
 51 DNHKADISSM FNEALIDIDS IKNAGRVF

11AA_SEQUENCE 1.0
 ID R63602 standard; Peptide; 367 AA.
 AC R63602:
 DT 09-MAY-1995 (first entry)
 DE MAP-kinase-phosphatase Ctl100.
 KW MAP-kinase-inhibitor; MAPK; MAP-kinase-phosphatase; Ctl100; STY8;
 RW neurological disease; proliferative disorder; inflammation;
 OS cardiovascular disease.
 PS Homo sapiens.
 PN MO9423039-A.
 PD 13-OCT-1994.
 PE 31-MAR-1994; G00694.
 PF 07-MAR-1993; GB-002750.
 PR 10-FEB-1994; GB-002573.
 PA (CANC-) CANCER RES INST.
 PI (CANC-) INST CANCER RES ROYAL CANCER HOSPITAL.
 DR Ashworth A, Hughes DA, Marshall CJ;
 PI WPI: 94-333199/41.
 PT Screening for a substance which is an inhibitor of mammalian MAP
 kinase - and is used in the mfr. of a medicament for the
 PT treatment of e.g. an inflammatory disorder or neurological
 PT disease.
 PS disclosure: Fig. 7d: 53pp; English.
 CC STY8 cDNA (given in 072864), encoding a peptide having
 CC MAP-kinase-phosphatase activity, was derived from a human
 CC brain cDNA library. The STY8 protein (R63601) encoded by this
 CC full-length cDNA showed homology to the previously isolated
 CC human Ctl100 gene product (R63602).
 SQ Sequence 367 AA.

R63602 Length: 367 February 11, 2000 15:48 Type: P Check: 5661

1 MNEVETLDA GGLRALIGER AAOCILLDR SFENFNCHI AGSVNFRST
 51 IYRRRAKGM GLEHIVPNAE LRGRLLAGAV HAVILIDERS AALDAKRRDG
 101 TLALAAFALC REARAAYVF LKGYEAFSA SCPELCSKOS TPMGLSLPLS
 151 TSVPDSAESG CSSCSTPLYD QGGPVEILPE IYLSAVYAS RKMDLALGI
 201 TALINVSANC PNHEGHYQY KSIPVEDNKH ADISSENEEA IDPDISIKNA
 251 GGRFVHCOA GISRSATICL AYLMRTNRYK IDEAFEFKQ RRSIIISPNS
 301 FMGQLLOFES QVLPHSCSAE AGSPAAVLD RGTSTTVFN FVPSIPVHST
 351 NSALSTYHSP ITTSPSC

11AA_SEQUENCE 1.0
 ID R58635 standard; protein; 124 AA.
 AC R58635:
 DT 15-MAY-1995 (first entry)
 DE Amylase inhibitor protein 0.26 AI.
 KW Amylase inhibitor; wheat; cation exchange; inhibitor; amylase; human;
 RW pancreatic; alpha-amylase; homodimer; blood glucose level;
 KM insulin secretion; appetite; food additive.
 OS Triticum durum.
 PN EP-618228-A.
 PD 05-OCT-1994; 302148.
 PE 24-MAR-1994; 302148.
 PF 28-MAR-1993; DE-091881.
 PR 28-MAY-1993; JP-148423.
 PA (NAGA-) NAGATA SANGYO CO. LTD.
 PI (NISS) NISSHIN FLOUR MILLING CO.
 DR Matsubara H, Miyazaki T, Morimoto T, Murayama R;
 PI WPI: 94-304407/38.
 PT New amylase inhibitor and process of prepn. from wheat - is
 PT useful for inhibiting increase in blood glucose level,
 PT controlling insulin secretion, suppressing appetite and as a food
 PT additive
 PS Claim 1: Page 15; 21pp; English.
 CC The amino acid sequence of a novel protein from the amylase
 CC inhibitor (AI) family. By size and electrophoretic measurements,
 CC this protein has been called 0.26 AI. The protein was isolated from
 CC wheat extracts by a novel process comprising (i) extracting a
 CC wheat-based material to produce a solution containing the AI, (ii)
 CC optionally purifying the solution to remove contaminants using a cation
 CC exchanger to absorb the AI, (iii) treating this exchanger with alkali
 CC at pH 9-13 to elute the AI, (iv) immediately adjusting the pH of the
 CC elute within neutral or acid range and (v) recovering the AI from this
 CC solution. The novel AI has a very high inhibitory activity against an
 CC amylase especially human pancreatic alpha-amylase. The activity of
 CC the AI is 5-35 times higher than others amylases of wheat origin such as
 CC 0.33 AI or 0.28 AI. The protein occurs as a homodimer with subunits of
 CC molecular masses 12500. The protein can be used in combination with
 CC another amylase inhibitor (0.19) (R58636) as an agent for inhibiting
 CC increase in blood glucose level, controlling insulin secretion or
 CC suppressing appetite, and as a food additive, the total content of both
 CC proteins in the material being at least 20% by weight.
 SQ Sequence 124 AA.

R58635 Length: 124 February 11, 2000 15:48 Type: P Check: 2685

1 SGPMWCYPGY ARKVPALPC RPVLKIQNG SOVPEAVLBD CCQGLADISE
 51 WCRGALYSM LPSMKYKGV QEGDAGTGAF PSRCREVVL TVASITAVCK
 101 LPIVIDASGD GAVYCKGVAA YPDA

11AA_SEQUENCE 1.0
 ID R58636 standard; protein; 124 AA.
 AC R58636:

DT 15-MAY-1995 (first entry)
 DE Amylase inhibitor protein 0.19 AI.
 KW Amylase inhibitor; wheat; cation exchange; inhibitor; amylase; human;
 KW pancreatic; alpha-amylase; homodimer; blood glucose level;
 OS Insulin secretion; appetite; food additive.
 PN Triticum durum.
 PD EP-618229-A.
 PD 05-OCT-1994.
 PD 24-MAR-1994; 302148.
 PR 29-MAR-1993; JP-091881.
 PR 28-MAY-1993; JP-148423.
 PA (NAGA-) NAGATA SANGYO CO LTD.
 PA (NISS) NISSHIN FLOUR MILLING CO.
 PI Matsubara H, Miyazaki T, Morimoto T, Murayama R;
 DR WPI; 94-304407/38.
 PT New amylase inhibitor and process of prep. from wheat - 1s
 PT useful for inhibiting increase in blood glucose level.
 PT controlling insulin secretion, suppressing appetite and as a food
 PT additive
 PS Claim 6: Page 16: 21pp; English.
 CC The amino acid sequence of a protein from the amylase inhibitor (AI)
 CC family. By size and electrophoretic measurements, this protein is called
 CC 0.19 AI. The protein occurs as a homodimer, the subunits having the amino
 CC acid sequence given. This protein can be used in combination with a
 CC novel amylase inhibitor (0.26) (R58635) as an agent for inhibiting
 CC increase in blood glucose level, controlling insulin secretion or
 CC suppressing appetite, and as a food additive, the total content of both
 CC proteins in the material being at least 20% by weight. The novel protein
 CC was isolated from wheat extracts by a novel process comprising
 CC (1) extracting a wheat-based material to produce a solution containing
 CC the AI, (11) optionally purifying the solution to remove contaminants
 CC using a cation exchanger to absorb the AI, (111) treating this exchanger
 CC with alkali at pH 9-13 to elute the AI, (1v) immediately adjusting the pH
 CC of the elute within neutral or acid range and (V) recovering the AI from
 CC this solution. The novel AI has a very high inhibitory activity against
 CC an amylase especially human pancreatic alpha-amylase. The activity of
 CC the AI is 5-30 times higher than other amylases of wheat origin such as
 CC 0.53 AI or 0.28 AI.
 SQ Sequence 124 AA;
 R58636 Length: 124 February 11, 2000 15:48 Type: P Check: 3320 ..
 1 SGPWMCYPGQ AFQVPALPAC RPLRLRQNG SOVEPAVLAD CCOQLAHISE
 51 WRCGALYSM LDMYKXGHA QEGQAGTGAF PROCRREVYKL TASTAVCR
 101 LPYVDAGSD GAYVCKRVAA YPDA
 !!AA-SEQUENCE 1.0
 ID R58587 standard; Protein; 246 AA.
 AC R58587;
 DT 09-MAY-1995 (first entry)
 DE Nicotinamide adenine dinucleotide synthetase N-terminal fragment.
 KW Nicotinamide adenine dinucleotide; NAD; synthetase; NADS; E.coli.
 OS Bacillus stearothermophilus.
 FT Key Location/Qualifiers
 FT Peptide 2..245
 FT /label= mature peptide
 FT J06225771-A.
 FN 16-AUG-1994.
 PD 01-FEB-1993; 014650.
 PD 01-FEB-1993; JP-014650.
 PA (ASAH) ASAH KASEI KOGYO KK.
 DR WPI; 94-298801/37.
 DR N-PSDB; 070537.
 PT DNA coding NAD synthetase - used for the prepn. of NAD synthetase
 PS Claim 2: Column 17-20; 16pp; Japanese.
 CC The amino acid sequence for the N-terminal region of nicotinamide
 CC adenine dinucleotide (NAD) synthetase (NADS). The gene encodes a protein
 CC of 245 amino acids. The gene was inserted into E.coli in an expression
 CC plasmid. This was used to produce the NADS protein. This method allows
 CC the efficient preparation of NADS polypeptide.
 SQ Sequence 246 AA;

R58587 Length: 246 February 11, 2000 15:48 Type: P Check: 9708 ..
 1 MQEIKDLVQ WLRDQSSAG LNGAVGISC GIDSAYVAHL IKRAFPPDSL
 51 GLIIPCKSNP KMEEDALKVY KSCGIRHLVT DTEARHLF GAVAEKAI
 101 GENSEERARL GDANTRARLR MTLIYAVANN YGILVAGTDN AAEMHTGYFT
 151 KYDGDGVLDV PLIHFTKGEV REMGRLLGVP EEIKKAPSA GLMEQGTDES
 201 EMGTYEMID KYLKGEEIPE RDRKIETRLH ERSHHKROLA IAPPEF
 !!AA-SEQUENCE 1.0
 ID R63443 standard; Protein; 120 AA.
 AC R63443;
 DT 06-JUN-1995 (first entry)
 DE Amino terminal sequence of the rubisco activase - AGP small subunit
 DE fusion enzyme.
 KW Barley; endosperm; ADP glucose pyrophosphorylase; AGP; starch;
 KW enzyme; transgenic plant; transit peptide; spinach; pps84; ss.
 OS Synthetic.
 PN WO9424292-A.
 PD 27-OCT-1994.
 PD 07-APR-1994; E01082.
 PR 08-APR-1993; GB-007408.
 PA (DANI-) DANISCO AS
 PI Kleczkowski L, Marcussen J, Okkels F, Olsen O, Poulsen P;
 DR WPI; 94-341871/42.
 PT Transgenic starch producing organism - comprises exogenous ADP
 PT glucose pyrophosphorylase enzyme DNA sequence
 PS Example; Fig 8; 88pp; English.
 CC Plasmid pps84 is a pVicoTriv deriv. in which a 3 kb KpnI fragment
 CC coding the construct:
 CC patatin promoter - spinach rubisco activase transit peptide -
 CC small subunit AGP from barley endosperm - 35S terminator;
 CC is inserted in the KpnI site. R63443 gives the N-terminal sequence
 CC of the rubisco activase-AGP small subunit fusion enzyme. pps84
 CC is claimed.
 SQ Sequence 120 AA;
 R63443 Length: 120 February 11, 2000 15:48 Type: P Check: 8722 ..
 1 MATASTVGA ATRAPLNG SSAGASVPTS GFGSSLSKH TVRPPSSSR
 51 TTSMTVAAE NEENKTDKMA HIAKDFSDQ LDIRRGKGVY DSLGMDYPL
 101 ASKVPLEPSS KHEQCNVYSH
 !!AA-SEQUENCE 1.0
 ID R63444 standard; Protein; 120 AA.
 AC R63444;
 DT 06-JUN-1995 (first entry)
 DE Amino terminal sequence of the rubisco activase - AGP large subunit
 DE fusion enzyme.
 KW Barley; endosperm; ADP glucose pyrophosphorylase; AGP; starch;
 KW enzyme; transgenic plant; transit peptide; spinach; pps14; ss.
 OS Synthetic.
 PN WO9424292-A.
 PD 27-OCT-1994.
 PD 07-APR-1994; E01082.
 PR 08-APR-1993; GB-007408.
 PA (DANI-) DANISCO AS
 PI Kleczkowski L, Marcussen J, Okkels F, Olsen O, Poulsen P;
 DR WPI; 94-341871/42.
 PT Transgenic starch producing organism - comprises exogenous ADP
 PT glucose pyrophosphorylase enzyme DNA sequence
 PS Example; Fig 8; 88pp; English.
 CC Plasmid pps14 is a pVicoTriv SGIN Man deriv. in which a 3.2 kb
 CC EcorI fragment contg. the construct:
 CC patatin promoter - spinach rubisco activase transit peptide -

CC large subunit AGP from barley endosperm - 355 terminator;
 CC is inserted in the EcoRI site. R63444 gives the N-terminal sequence
 CC of the rubisco activase-AGP large subunit fusion enzyme. pep14
 CC is claimed.
 SQ Sequence 120 AA;

R63444 Length: 120 February 11, 2000 15:48 Type: P Check: 7811 ..

1 MATAVSTVGA ATRAPPLNG SSAGSVPTS GFLGSLSKKH TNRPPSSSR
 51 TTSMTVKAEE NEEKNDKMA HIAKDFSDQ LDIRRGKGV DSLGIMQFS
 101 SVLPLEGKAC VSPVREGSA

!!AA_SEQUENCE 1.0
 ID R62507 standard; Protein; 313 AA.
 AC R62507;
 DT 26-JUN-1995 (first entry)
 DE Galactosyl transferase 3' clone product.
 KW Gal-alpha (1.3) galactosyl transferase; xenograft; transplant;
 KM rejection.
 OS Sus scrofa domestica.
 PN WO9421799-A.
 PD 29-SEP-1994.
 PR 15-MAR-1994; AU0126.
 PR 16-MAR-1993; AU-007854.
 PA (AUST-) AUSTIN RES INST.
 PI McKenzie IFC, Sandrin MS;
 DR N-PSDB: 074711.
 DR N-PSDB: 94-317019/39.
 PT DNA sequences encoding gal-alpha (1.3) galactosyl transferase -
 PT and clones contg. such sequences are used in xenograft therapies
 PS The sequence is that of the product of the porcine gal-alpha (1.3)
 CC galactosyl transferase gene which produces a Gal epitope on the
 CC surface of porcine cells. This epitope is recognised by antibodies
 CC which are responsible for hyperacute rejection of xenotransplanted
 CC pig cells, tissues and organs.
 CC See also R62508.
 CC Sequence 313 AA;
 SQ

R62507 Length: 313 February 11, 2000 15:48 Type: P Check: 2303 ..

1 VPSSNSASQS PQAMDPDPCSP RLSTLSKAIL TLGFTVRKPP EVVITRMKA
 51 PVWEGTYNR AVIDNTYAKQ KITVGLVFA VGRYIEHYLE ELISANTYF
 101 MVGRKVIETI MDDDISRMPL IELGPLNSFK VEFIKSEKRV QDISMAMKMT
 151 IGEHLIAHQ HEVDLEFICID VDQVFQNNFG VETLGQSVAK QAMWYIAHP
 201 DEFTYERKE SAAVIFPRQG DYYHAALIG GPPTQVNLIT QCEKGIQID
 251 KENDIEAEWH DESGLNKYFL INKPTKLISL EYCMYHIGM SYDIRIYVGA
 301 WQREYNLYR NMI

!!AA_SEQUENCE 1.0
 ID R63355 standard; peptide; 216 AA.
 AC R63355;
 DT 28-JUN-1995 (first entry)
 DE Ii protein.
 KW Ii protein; immunomodulator; antigen presentation; MHC;
 KW major histocompatibility complex.
 OS Homo sapiens.
 PN WO9426773-A.
 PD 24-NOV-1994.
 PR 18-MAY-1994; U05617.
 PR 19-MAY-1993; US-064400.
 PA (ANTI-) ANTIGEN EXPRESS INC.
 PI Humphreys RE;
 DR WPI: 95-006697/01.
 PT Identification of mutant Ii mols having altered endoprotease

PR cleavage sites - and useful for identifying immunomodulators,
 PR esp. peptide(s) which enhance or inhibit antigen presentation
 PS Disclosed: page 38-39: 50pp; English.
 CC Immunomodulatory peptides of Ii protein (full sequence given
 CC R63355) have been identified. The Ii peptide given in R63352
 CC enhances MHC class II-mediated antigen presentation, while the
 CC Ii peptides given in R63353-54 inhibit antigen presentation.
 CC Sequence 216 AA;
 SQ

R63355 Length: 216 February 11, 2000 15:48 Type: P Check: 1046 ..

1 MDDQDLISN NEQPLMKRR POAPESCSR GALTYGFSIL VTLLAGQAT
 51 TAYFLYQOQG RLDKLTFTSQ NLOENLRMK LPPRPKPVSK MRMATPLMQ
 101 ALPMGALPQG PMQNAIRYGN MTEHVMHL QNADPLKAYP PLKGFPENL
 151 RLKNTMETL DWKVESMMH HWLLEMSRH SLEQRTDAP PKESIELED
 201 SSGLGVTROD LGPVPM

!!AA_SEQUENCE 1.0
 ID R63159 standard; Protein; 376 AA.
 AC R63159;
 DT 23-JUN-1995 (first entry)
 DE Mouse growth differentiation factor-8 protein.
 KW Growth differentiation factor-8; GDF-8; cell proliferation;
 KM adipocyte; obesity; transforming growth factor-beta.
 OS Mus musculus.
 PN WO9421681-A.
 PD 29-SEP-1994.
 PR 18-MAR-1994; U03019.
 PR 19-MAR-1993; US-033923.
 PA (UYXO) UNIV JOHNS HOPKINS SCHOOL MED.
 PI Ise S, McPherron AC;
 DR WPI: 94-316943/39.
 DR Q-PSDB: 076371.
 PT New growth differentiation factor 8 - useful for treatment and
 PT diagnosis of cell proliferative disorders esp. of muscle.
 PS Claim 3; page 47: 84pp; English.
 CC GDF-8 can be used to maintain cells before transplantation; to
 CC improve efficiency of cell fusion and to treat obesity or diseases
 CC related to abnormal adipocyte proliferation.
 CC Sequence 376 AA;
 SQ

R63159 Length: 376 February 11, 2000 15:48 Type: P Check: 2293 ..

1 MMOKLQWYV IYLEMLIAG PVLDNBSER EENVKEKELC NACAMQNTR
 51 YSRIEAIRIQ ILSKLELETA PNISKDAIRO LLPRAPPLRE LIDQYDQRD
 101 DSDSLELED DYHATETII TMTESDFLM QADGPKCCF FKSSKIQYN
 151 KYVRAQLMIY LREVKPTTV FVOILRLIKP MKDGRIRYGI RSLKDKSPG
 201 TGIMQSIDVK TVLQNLAKOP ESNGLIEIKA LDENGHOLAV TFPQSGDGL
 251 NFLEVKVTD TPKRSRRDFG LDODEHSTES RCRRIPLVVD FEAGNMWII
 301 AFRYKANYC SGCECFVFLQ KYPHILVHQ ANRGSAGPC CTPIKSPIN
 351 MLYENGKEOI IYKIPAMVY DRGCS

!!AA_SEQUENCE 1.0
 ID R63160 standard; Protein; 375 AA.
 AC R63160;
 DT 23-JUN-1995 (first entry)
 DE Human growth differentiation factor-8 protein.
 KW Growth differentiation factor-8; GDF-8; cell proliferation;
 KM adipocyte; obesity; transforming growth factor-beta.
 OS Homo sapiens.
 PN WO9421681-A.
 PD 29-SEP-1994.

PF 18-MAR-1994; U03019.
 PR 19-MAR-1993; US-033923.
 PA (UYUO) UNIV. JOHNS HOPKINS SCHOOL MED.
 PI Lee S, McPherson AC;
 DR WPI; 94-316943/39.
 DR O-PSDB; Q76372.
 PT New growth differentiation factor 8 - useful for treatment and
 PT diagnosis of cell proliferative disorders esp. of muscle.
 PS Claim 3; Page 58; 84pp; English.
 CC GDF-8 can be used to maintain cells before transplantation; to
 CC improve efficiency of cell fusion and to treat obesity or diseases
 CC related to abnormal adipocyte proliferation.
 SQ Sequence 375 AA;

R63160 Length: 375 February 11, 2000 15:48 Type: P Check: 1814 ..

1 MQLQGCYVI YLFMLIVAGP VDLNSENSEK ENVEKEGICN ACTWRQNTKS
 51 SREAKIKOI LSKRLERAP NISKQVIRQL LPAKPLREL IDQYDVGIRD
 101 SSGSLEDDE YHATETITIT MPTESDPLAQ VDGKRCCKP KFSKTIYNK
 151 YKAQIMITL RPVEPTTIV VQILRLIKPM KGTGRTGIR SLKIDMNPCT
 201 GIMQSIDVKT VLQNMKQPE SNLGEIKAL DENGHLAVT PPGEGEDGLN
 251 PLELVKVTDT PKRSRDFGL DCDHSTESR CCRYPITVDF EAFGMDIITA
 301 PKRYKANYCS GECEFEVLQK YPTHVLVQA NPGSGAGPC TPTKSPDIN
 351 LYFNGKEQII YGKIPAMVVD RCGCS

!!AA_SEQUENCE 1.0
 ID R6665 standard; peptide; 25 AA.
 AC R6665;
 DT 21-JUL-1995 (first entry)
 DE Binding peptide B (84 TS)
 KW Inner capsid protein; VP6; carrier; vaccine; rotavirus; VP3;
 KM binding peptide.
 OS Rotavirus
 PN US5374426-A.
 PD 20-DEC-1994.
 PF 03-SEP-1986; 903222.
 PR 03-SEP-1986; US-903222.
 PR 02-SEP-1987; US-092120.
 PR 30-OCT-1989; US-429147.
 PA (UYSA-) UNIV. SASKATCHEWAN.
 PI Frenchick PJ, Mullin-ready KF, Sabara MI;
 DR WPI; 95-035623/05.
 PT Immunogenic complex contg. rotavirus VP6 protein as carrier -
 PT coupled, esp. by protein-protein interaction, to an epitope
 PT contg. molecule.
 PS Claim 11(b); Column 26; 23pp; English.
 CC An immunological carrier complex comprises an epitope-bearing
 CC molecule bound to the rotavirus inner capsid protein VP6 (given in
 CC R66669) using a binding peptide such as those given in R66664-8.
 CC Peptide B (or 84 TS) (R66665) is derived from rotavirus VP3
 CC (231-254).
 SQ Sequence 25 AA;

R66665 Length: 25 February 11, 2000 15:48 Type: P Check: 4995 ..

1 CNIPASIVS RNIVYTRAOP NODIA

!!AA_SEQUENCE 1.0
 ID R66680 standard; peptide; 25 AA.
 AC R66680;
 DT 21-JUL-1995 (first entry)
 DE Peptide 84 TS-CYS.
 KW Inner capsid protein; VP6; carrier; vaccine; rotavirus; VP3;
 KM binding peptide.
 OS Synthetic.
 PN US5374426-A.
 PR 28-OCT-1993; DK-001217.

PD 20-DEC-1994.
 PF 03-SEP-1986; 903222.
 PR 03-SEP-1986; US-903222.
 PR 02-SEP-1987; US-092120.
 PR 30-OCT-1989; US-429147.
 PA (UYSA-) UNIV. SASKATCHEWAN.
 PI Frenchick PJ, Mullin-ready KF, Sabara MI;
 DR WPI; 95-035623/05.
 PT Immunogenic complex contg. rotavirus VP6 protein as carrier -
 PT coupled, esp. by protein-protein interaction, to an epitope
 PT contg. molecule.
 PS Example 2; Column 18; 23pp; English.
 CC An immunological carrier complex comprises an epitope-bearing
 CC molecule bound to the rotavirus inner capsid protein VP6 (given in
 CC R66669) using a binding peptide such as peptide B (84 TS)
 CC (R66665). Variants of peptide B were also tested for binding to VP6.
 CC Only those that generated a spatial arrangement of Cys and Arg
 CC residues in a suitable 3-dimensional orientation (e.g. R66666-68)
 CC were suitable. 84 TS-CYS (R66680) did not bind to VP6.
 SQ Sequence 25 AA;

R66680 Length: 25 February 11, 2000 15:48 Type: P Check: 5136 ..

1 NISPASIVSR NIVYTRAOPN OSIAC

!!AA_SEQUENCE 1.0
 ID R66682 standard; peptide; 25 AA.
 AC R66682;
 DT 21-JUL-1995 (first entry)
 DE Peptide MONOSER.
 KW Inner capsid protein; VP6; carrier; vaccine; rotavirus; VP3;
 KM binding peptide.
 OS Synthetic.
 PN US5374426-A.
 PD 20-DEC-1994.
 PF 03-SEP-1986; 903222.
 PR 03-SEP-1986; US-903222.
 PR 02-SEP-1987; US-092120.
 PR 30-OCT-1989; US-429147.
 PA (UYSA-) UNIV. SASKATCHEWAN.
 PI Frenchick PJ, Mullin-ready KF, Sabara MI;
 DR WPI; 95-035623/05.
 PT Immunogenic complex contg. rotavirus VP6 protein as carrier -
 PT coupled, esp. by protein-protein interaction, to an epitope
 PT contg. molecule.
 PS Example 2; Column 18; 23pp; English.
 CC An immunological carrier complex comprises an epitope-bearing
 CC molecule bound to the rotavirus inner capsid protein VP6 (given in
 CC R66669) using a binding peptide such as peptide B (84 TS)
 CC (R66665). Variants of peptide B were also tested for binding to VP6.
 CC Only those that generated a spatial arrangement of Cys and Arg
 CC residues in a suitable 3-dimensional orientation (e.g. R66666-68)
 CC were suitable. MONOSER (R66682) did not bind to VP6.
 SQ Sequence 25 AA;

R66682 Length: 25 February 11, 2000 15:48 Type: P Check: 5012 ..

1 CNIPASIVS RNIVYTSAPQ NODIA

!!AA_SEQUENCE 1.0
 ID R67596 standard; Protein; 331 AA.
 AC R67596;
 DT 20-JUL-1995 (first entry)
 DE A. aculeatus pectin methyltransferase.
 KW Pectin methyltransferase; PME; strain CBS 101.43; demethylation; pectin;
 KW food industry; fruit; vegetable; processing; plant cell wall;
 KW modification; pectin lyase; glucanase; firmness.
 OS Aspergillus aculeatus.
 PN WO9425575-A.
 PD 10-NOV-1994.
 PD 10-NOV-1994.
 PF 28-APR-1994; DK0173.
 PR 30-APR-1993; DK-000487.
 PR 28-OCT-1993; DK-001217.

(NOVO) NOVO-NORDISK AS.
 PA Andersen LN, Budolfesen G, Christgau S, Dalboege H;
 PI Heide-Hansen HP, Kauppinen S, Kofoed LV;
 DR WPI: 95-022253/03.
 DR N-PSDB: 079154.
 PT Enzyme with pectin methyltransferase activity - useful for
 PT demethylation of pectin, e.g. for application in food industry.
 PS Claim 1, Page 37, 53pp; English.
 CC This sequence represents an enzyme showing pectin methyltransferase (PME)
 CC activity. The DNA sequence is useful for demethylation of
 CC strain CGS 101.43. The PME enzyme is useful for demethylation of
 CC pectin. It can therefore be used in the food industry for fruit and
 CC vegetable processing; it modifies plant cell walls and can be used
 CC with other plant cell wall degrading enzymes, eg. pectin lyase,
 CC glucanase, etc. It can be used to improve the firmness of a pectin-
 CC containing material or to increase viscosity of a pectin-contg. material.
 CC The enzyme has an optimum pH of 4.5, a temperature optimum of 45 deg
 CC C and a molecular weight of 43 kD.
 SQ Sequence 331 AA;
 R67596 Length: 331 February 11, 2000 15:48 Type: P Check: 5465 ..
 1 MKKSVLASAT FAVSATAASR TPAPGATV AKSGGYTTI GDAIDALSTS
 51 TTDITQITIE EGYDEQVYL PMTKGVITII GQIENDISVA DMTVITTHAI
 101 SYEDAGESDD TTAIFRNKAV GSOVYNTNIA NTCGQACHQA LALSMAAOQ
 151 GYGCNFTGY QDTTAQTGN QYINSYIEG AVDEFICQHA RAMFQNDIR
 201 VEGPTASASI TANGRSSETD TSYVYINNST VAAKESDVA EGYTTGRPM
 251 SEFARVFEQO TSMNTVINST GMTWSTSTP NTEVTFGEY ANOPAPAPRA
 301 PAHVAEKTDK KTTITDITGS DYTSMVDISY F
 11AA_SEQUENCE 1.0
 ID R71324 standard; Protein: 329 AA.
 AC R71324;
 DE 21-OCT-1995 (first entry)
 DE Acetyl-CoA-reductase; transgenic plant;
 KW Acetyl-CoA-hydroxylalkanoate: poly-beta-hydroxybutyrate;
 KW biodegradable thermoplastic.
 OS Alcaligenes eutrophus.
 PN WO9505472-A.
 PD 23-FEB-1995.
 PF 17-AUG-1994; 009265.
 PR 17-AUG-1993; US-108193.
 PR 06-JUN-1994; US-254357.
 PA (UNMS) UNIV MICHIGAN STATE.
 PI Nawrath C, Pollier Y, Somerville CR;
 DR WPI: 95-098770/13.
 DR N-PSDB: 085642.
 PT Transgenic plant material with plastid(s) contg. the enzymes for
 PT synthesis of poly(hydroxyalkanoate(s)) - express
 PT poly(hydroxybutyrate and have good growth and seed formation.
 PS Claim 2; Page 60-61; 88pp; English.
 CC The acetyl-CoA-reductase gene (phbB) from A. eutrophus is cloned
 CC under the control of an Arabidopsis thaliana seed storage protein
 CC promoter for plastid tissue-specific gene expression in a
 CC transgenic plant. When expressed with the 3-ketochiolase
 CC (phbA) and poly-beta-hydroxyalkanoate-synthase (phbC) genes, a
 CC poly-beta-hydroxyalkanoate (PHA), specifically poly-beta-
 CC hydroxybutyrate (PHB), is expressed in the transgenic plant
 CC (preferably a Brassica e.g. rape). PHB and related PHAs are
 CC biodegradable thermoplastics with many useful applications.
 SQ Sequence 329 AA;
 R71324 Length: 329 February 11, 2000 15:48 Type: P Check: 7564 ..
 1 MASMISNNV TTVSRASRGQ SAAVAPFGL KSMTGFPVK VNTDITSTS
 51 NGRVACMOV WPIQKKKE TLTSLPELR DSRVOTRIAY VTGNGGICIT
 101 AICORLAKOG FRVAVGCCPN SPRERKLEO QKALGFDFIA SEGNAVADMS
 151 TKTAFDKVS EYGEVDVLIN NAGITRDVVF RMTRADMDA VIDTNTLSLF
 201 NTKQVODGM ADRGGRILVN ISSVNGCKGQ FGOITVSTAK AGLAGHTMAL
 251 AOEATATGV VNTVSPGYIA TDWKAIRQD VLDKIVATIP VKRLGDPBEI
 301 ASICAMLSSE ESGFSTGDF SLNGLHMG
 11AA_SEQUENCE 1.0
 ID R70100 standard; Peptide: 274 AA.
 AC R70100; (first entry)
 DE Lettuce infectious yellows virus open reading frame 3 protein.
 DE Lettuce infectious yellows virus expression cassette; Agrobacterium tumefaciens; plant;
 KW Expression vector; expression cassette; Agrobacterium tumefaciens; plant;
 KW Lettuce infectious yellows virus coat protein; amplify; PCR; primer;
 KW heat shock protein.
 OS Synthetic.
 PN WO9502056-A.
 PD 19-JAN-1995.
 PF 13-JUN-1994; 006430.
 PF 09-JUL-1993; US-090532.
 PR 15-OCT-1993; US-138138.
 PR 01-NOV-1993; US-146780.
 PA (REGC) UNIV CALIFORNIA.
 PA (UPJO) UPJOHN CO.
 PI Boeshore ML, Carney KJ, Falk BW, Klaassen V;
 DR WPI: 95-106554/14.
 DR N-PSDB: 083174.
 PT New DNA encoding proteins of lettuce infectious yellows virus
 PT and related vectors, transformed cells and transgenic plants
 PT resistant to virus infection
 PS Claim 1; Fig 14; 92pp; English.
 CC The amino acid sequence encoded by the open reading frame 3 of the
 CC lettuce infectious yellows virus (LIYV) RNA 1. The gene was amplified by
 CC PCR using the primers 083175-6. The resultant product was used to
 CC construct expression cassettes expressing the gene as a sense (pEP6170)
 CC message. The vectors containing the cassettes can be transformed into
 CC Agrobacterium tumefaciens for transfer of the genes into plants.
 CC containing the sense or antisense vectors are resistant to LIYV
 CC infections.
 SQ Sequence 274 AA;
 R70100 Length: 274 February 11, 2000 15:48 Type: P Check: 4652 ..
 1 MIMMSPIYAL TKOCYIDTAY RLAVPQCHQA IYVACRIIE LSYGEWITV
 51 LCGFKMDTSS FIASIEKDNL MCCLISLVEM RDRRLCNDP PILNGVNL
 101 ELLIGKRLNK INNLKNCYVI RELITINISK EWSGQALKV GLHCFNLSSQ
 151 AESRRVYLL SDKESLNKRN FSRIYVPKVY TDLIYDLIGV LYVNTGVNID
 201 LVEKFIIDKL EFLVYDGEQ FKSPOVEYND ICTVYNLKPI IKYRNHTDQ
 251 SIVIECDVI GKGINKTKK ICKN
 11AA_SEQUENCE 1.0
 ID R72715 standard; Protein: 359 AA.
 AC R72715;
 DE 02-DEC-1995 (first entry)
 DE hisC gene product of Salmonella typhimurium.
 KW hisG; hisC; Salmonella typhimurium; mutant; auxotroph; detection;
 KW identification; mutagen; transition; transversion.
 OS Salmonella typhimurium.
 PN WO9510629-A.
 PD 20-APR-1995.
 PF 14-OCT-1994; U11617.
 PF 15-OCT-1993; US-137627.
 PA (REGC) UNIV CALIFORNIA.

PI Ames BN, Gee P, Maron DM;
 DR WPI: 95-161813/21.
 DR N-PSDB: 089706.
 PT *Salmonella* auxotrophic mutants for detecting mutagenic agents -
 PT Disclosure; Figure 4; 6pp; English.
 CC Analogues of Type I RIP are defined as non-naturally occurring
 CC polypeptides that share the ribosome-inactivating activity of the
 CC natural protein but differ in AA sequence. Preferred analogues have
 CC a Cys available for disulfide bonding located at a posn. It is AA
 CC sequence from the posn. corresp. to posn. 251 in ricin A-chain RTA
 CC to the carboxy terminus of the analogue. (R74176 is the sequence of
 CC ricin A-chain RTA, which is a Type II RIP). The primary AA
 CC sequence of the Type I RIPS gelonin, BRP, momordin II, luffin
 CC (see Islam et al. Agricultural Biological Chem., 54(5) 1343-45 1991),
 CC (1980), momordin I (see Chow et al., J. Biol. Chem., 265, 8670-74
 CC 1990), momordin I (see Ho et al., BBA, 1088, 311-14 1991),
 CC Mirabilis antiviral protein (see Habuku et al., J. Biol. Chem.,
 CC 264(12) 6629-37 1989), pokeweed antiviral protein isolated from
 CC seeds (see Kung et al., Agric. Biol. Chem., 54(12), 3301-18 1990)
 CC and seporin (see Banerji et al., Eur. J. Biochem., 183, 465-70
 CC 1989) are individually aligned with the primary sequence of the
 CC ricin A-chain (see Halling et al., Nucleic Acids Res., 13,
 CC 8019-8033 1985) respectively in Figures 1-9. The AAs invariant
 CC among the ricin A-chain and the Type I RIPS are indicated in FT.
 SQ Sequence 359 AA;
 R72715 Length: 359 February 11, 2000 15:48 Type: P Check: 516 ..
 1 MSTEHTLSVA DLARENVRNL VPQSAARRIG GNGDWMINAN EEPFAVEFOL
 51 TQOTLNRYPE COPKAVIENY AQYAGVKPEO VLVSRADEG IELVIRAFCE
 101 PKDAILYCP PTYGMYSVA ETIGVERRTV PALENMQLDL QGISDMLDGT
 151 KVFPCSPNN PTGQLNPQD LRTLELTRG KAIYVADEAY IERCPOATLT
 201 GMLVEYPLV ILRTLSKAPA LAGLRGFTL ANEEVINLL KYIAPPLST
 251 PVADIAAAL CPQGINAMRD RYAQYQERO YLVNALQQA CVENYDSET
 301 NTLARFTAS SSVFKSLMDQ GILLRDNKO PLSGCLRTT VGRROZNRV
 351 IDALRAEPV
 !!AA_SEQUENCE 1.0
 ID R74180 standard; protein: 248 AA.
 AC R74180;
 DT 01-JAN-1996 (first entry)
 DE Type I ribosome-inactivating protein luffin.
 KW Ribosome inactivating protein; RIP; Type I; cytotoxin; immunotoxin.
 OS Plant.
 FH Key Location/Qualifiers
 FT misc_difference 14 /label- invariant residue
 FT /note- "In ricin A-chain and the Type I RIPS"
 FT misc_difference 22 /label- see above
 FT misc_difference 70 /label- see above
 FT misc_difference 110 /label- see above
 FT misc_difference 131 /label- see above
 FT misc_difference 157. 158 /label- see above
 FT misc_difference 160 /label- see above
 FT misc_difference 192 /label- see above
 FT misc_difference 192 /label- see above
 PN US5416202-A.
 PD 16-MAY-1995.
 PF 09-DEC-1992; 988430.
 PR 04-NOV-1991; US-787567.
 PR 09-DEC-1992; US-988430.
 PA (XOMA) XOMA CORP.
 PI Bernhard SL, Better MD, Carroll SF, Lane JA, Lei SP;
 DR WPI: 95-193480/25.

PT Polynucleotide(s) encoding gelonin analogues - having a cystein
 PT residue for intermolecular bonding for the prodn. of immuno-toxins(s)
 PS Disclosure; Figure 4; 6pp; English.
 CC Analogues of Type I RIP are defined as non-naturally occurring
 CC polypeptides that share the ribosome-inactivating activity of the
 CC natural protein but differ in AA sequence. Preferred analogues have
 CC a Cys available for disulfide bonding located at a posn. It is AA
 CC sequence from the posn. corresp. to posn. 251 in ricin A-chain RTA
 CC to the carboxy terminus of the analogue. (R74176 is the sequence of
 CC ricin A-chain RTA, which is a Type II RIP). The primary AA
 CC sequence of the Type I RIPS gelonin, BRP, momordin II, luffin
 CC (see Islam et al. Agricultural Biological Chem., 54(5) 1343-45 1991),
 CC (1980), momordin I (see Chow et al., J. Biol. Chem., 265, 8670-74
 CC 1990), momordin I (see Ho et al., BBA, 1088, 311-14 1991),
 CC Mirabilis antiviral protein (see Habuku et al., J. Biol. Chem.,
 CC 264(12) 6629-37 1989), pokeweed antiviral protein isolated from
 CC seeds (see Kung et al., Agric. Biol. Chem., 54(12), 3301-18 1990)
 CC and seporin (see Banerji et al., Eur. J. Biochem., 183, 465-70
 CC 1989) are individually aligned with the primary sequence of the
 CC ricin A-chain (see Halling et al., Nucleic Acids Res., 13,
 CC 8019-8033 1985) respectively in Figures 1-9. The AAs invariant
 CC among the ricin A-chain and the Type I RIPS are indicated in FT.
 SQ Sequence 248 AA;
 R74180 Length: 248 February 11, 2000 15:48 Type: P Check: 8163 ..
 1 DVRSLSGSS STSYSKFTID LRKALPNSGT VYNLTLLSS ASGASRYTLM
 51 TLSNYDEKAI TAAVDVSOLY INGVLNSTS YFENESDARL ASQVFGNST
 101 IVTLPGSGY EKLQTAAGRI REKIPLEGPA LDGALTTHR YDSTAAAF
 151 LVILQTAEA SRFKYEGGI IERISKQVP SLATISLERS LMSLSKQIO
 201 LAQINNFTK TPVYITDDQO QREVEITVTS KVTAKNIQLL LNYQNVA
 !!AA_SEQUENCE 1.0
 ID R84103 standard; protein: 303 AA.
 AC R84103;
 DT 20-MAR-1996 (first entry)
 DE Equine herpesvirus 1 (EHV1) unique short (US2) gene prod.
 KW Equine herpesvirus 1; EHV1; unique short; US2; recombinant; vaccine;
 OS attenuated virus; diagnosis.
 FH Key Location/Qualifiers
 FT region 123..140 /note- "EHV-1, EHV-4, HSV-1, PRV, HSV-2, MDV, and
 FT /note- "IBR conserved US2 region"
 FT MW9522607-A1.
 PD 24-AUG-1995.
 PF 16-FEB-1995; U02087.
 PR 15-FEB-1994; US-198094.
 PA (SYTR) SYNTRO CORP.
 PI Chiang CH, Cochran MD;
 DR WPI: 95-302714/39.
 DR N-PSDB: T00531.
 DT Recombinant equine herpes viruses pref. contg. a deletion in a
 FT region not essential for replication - used in vaccines to protect
 FT horses from infection
 PS Example 11; Pages 86-87; 159pp; English.
 CC T00531 encodes R84103 the equine herpesvirus 1 (EHV) unique short 2
 CC (US2) gene prod. A recombinant attenuated EHV can be produced by
 CC deleting the above US2 gene (a region not essential for replication)
 CC from the viral DNA. The attenuated virus can be used as a foetal safe
 CC vaccine to protect an equine against EHV, or in a test to determine
 CC if an equine has been vaccinated against or is infected with EHV.
 SQ Sequence 303 AA;
 R84103 Length: 303 February 11, 2000 15:48 Type: P Check: 5723 ..
 1 MGVLITVTV VYDRKALPN SSIDVDGHLW EFLSRGCFVL ASPLIGITIV
 51 VRSDLYRFS SLLITPKAC RPIYTRGAT AIALDNGVY YHEDRMGVSI


```
FT /note="mouse heavy chain variable complementarity
FT determining region 2"
FT region
FT /label=.98
FT /label=FR3
FT /note="mouse heavy chain variable framework
FT region 3"
FT 98..112
FT /label=CDR3
FT /note="mouse heavy chain variable complementarity
FT determining region 3"
FT 113..123
FT /label=FR4
FT /note="mouse heavy light chain variable framework
FT region 4"
FT W09519790-A1.
FT 27-JUL-1995.
FT 25-JAN-1995: US-186269.
FT (AIHE-) ARHENA NEUROSCIENCES INC.
FT Bendig MM, Jones TS, Leger OJ, Saldanha J;
FT WPI: 95-269276/35.
FT New humanised antibodies against VLA-4 - used for inhibiting
FT leukocyte adhesion to endothelial cells, partic. for treating
FT inflammatory disease.
FT PS Disclosure: Page 68; 105pp; English.
FT The sequence represents the mouse anti-VLA-4 antibody 21.6 heavy chain
FT variable region (without signal sequence). Cloned cDNA CDR sequences of
FT mouse 21.6 variable light and variable heavy regions are linked to human
FT constant framework regions of the REI antibody for the light chain and
FT the 2CL antibody for the heavy chain in the construction of a humanized
FT antibody against VLA-4. The 5' and 3' ends of the mouse cDNAs are
FT modified using PCR primers (See Q99895-98) and then subcloned into
FT mammalian cell expression vectors containing human kappa or gamma-1,
FT constant regions. In the humanized light chain, amino acids L45, L49,
FT L58 and L69 in the human kappa LCVR framework are replaced by the amino
FT acid present in the equivalent position of the mouse 21.6 Ig light chain.
FT Plasmids encoding the chimeric antibodies are transfected into COS cells.
FT The humanized antibodies can be used to inhibit adhesion of a leukocyte
FT to an endothelial cell and to treat inflammatory diseases such as
FT multiple sclerosis. They can also be used in the treatment of stroke,
FT cerebral traumas, meningitis or encephalitis. The antibodies can also be
FT used for detecting VLA-4, for affinity purification or for generating
FT anti-idiotypic antibodies.
FT Sequence 123 AA;
R81330 Length: 123 February 11, 2000 15:48 Type: P Check: 8013
1 EVLOQSGAE LVKPGASVKL SCTASGENIK DVIHCVKOR PEOGLEWIGR
51 IDPANGYTRY DPKFGKATI TADTSSNTAY LQLSLTSBD TAYVFCAREG
101 YGNYGYVAM DYWGQSTVT VSS
IIAA_SEQUENCE 1.0
ID R81333 standard; Protein: 142 AA.
AC R81333:
DE 23-MAR-1996 (first entry)
DE Human VLA-4 reshaped antibody 21.6 light heavy variable region.
DE Humanized antibody; leukocyte adhesion molecule; VLA-4; therapeutic;
KW antibody engineering.
OS Homo sapiens.
DR Key
DR 1..19 Location/Qualifiers
FT peptide
FT /note="signal peptide"
FT 20..49
FT /note="framework region 1"
FT 50..54
FT /note="complementarity determining region 1"
FT 55..68
FT /note="framework region 2"
FT 69..85
FT /note="complementarity determining region 2"
FT 86..117
```

```
FT /note="framework region 3"
FT 118..131
FT /note="complementarity determining region 3"
FT 132..142
FT /note="framework region 4"
FT W09519790-A1.
FT 27-JUL-1995.
FT 25-JAN-1995: US-186269.
FT (AIHE-) ARHENA NEUROSCIENCES INC.
FT Bendig MM, Jones TS, Leger OJ, Saldanha J;
FT WPI: 95-269276/35.
FT N-PSDB: Q99894.
FT New humanised antibodies against VLA-4 - used for inhibiting
FT leukocyte adhesion to endothelial cells, partic. for treating
FT inflammatory disease.
FT PS Disclosure: Fig 11; 105pp; English.
FT The sequence represents the human reshaped antibody 21.6 heavy
FT chain variable region against leukocyte adhesion molecule VLA-4.
FT Cloned cDNA sequences of mouse 21.6 VH (Q99892) and VL (Q99889)
FT regions are linked to human constant regions in the construction
FT of a humanized antibody against VLA-4. The 5' and 3' ends of the
FT mouse cDNAs are modified using PCR primers (Q99895-98) and then
FT subcloned into mammalian cell expression vectors containing human
FT kappa or gamma-1 constant regions. In the humanized heavy chain,
FT amino acids H27, H28, H29, H30, H44 and H71 in the human HC VR
FT framework are replaced by the amino acid present in the equivalent
FT position of the mouse 21.6 Ig H chain. Plasmids encoding the
FT chimeric antibodies are transfected into COS cells. The humanized
FT antibodies can be used to inhibit adhesion of a leukocyte to an
FT endothelial cell and to treat inflammatory diseases such as multiple
FT sclerosis. They can also be used in the treatment of stroke,
FT cerebral traumas, meningitis or encephalitis. The antibodies can
FT also be used for detecting VLA-4, for affinity purification or for
FT generating anti-idiotypic antibodies.
FT Sequence 142 AA;
R81333 Length: 142 February 11, 2000 15:48 Type: P Check: 9019
1 MDWTRVEFL LAVAPGASHQ VOLVQGAEV KKPASVKS CRASGENIKD
51 TYIHWRAAP GQRLWMGRI DPANGTKYD PKFGRTVIT ADPSASTAYM
101 ELSLRSEDT AVYYCAREGY YGNYGYVAMD YMGGTLVTV SS
IIAA_SEQUENCE 1.0
ID R75202 standard; Protein: 225 AA.
AC R75202:
DE 10-MAY-1996 (first entry)
DE Tyrosine phosphatase MPTP-delta insertion sequence, INS-1.
DE Tyrosine phosphatase MPTP-delta; murine; brain tissue;
KW glutathione-S-transferase; fusion protein; E. coli; differentiation;
KW activation; information transmission; nervous system; immune system;
KW carcinogenesis; insertion; INS-1; INS-2.
OS Mus musculus domesticus.
DR 107236487-R.
DR 12-SEP-1995.
DR 28-FEB-1994: 054726.
DR 28-FEB-1994: JP-054726.
DR (TOKS-) TOKIOTO SHINKAI KAGAKU SOGO KENKYUSHO ZH.
DR WPI: 95-347455/45.
DR N-PSDB: Q94312.
DR DNA encoding tyrosine phosphatase MPTP delta - useful for
FT elucidation of signal transmission mechanisms.
FT Claim 2; Page 11-12; 14pp; Japanese.
FT The sequences given in R75202-03 are encoded by insertion sequences
FT which were included in the tyrosine phosphatase MPTP-delta coding
FT sequence. The INS-1 cDNA was inserted between the 66th A and the 67th
FT C of the MPTP-delta coding sequence and the INS-2 cDNA was inserted
FT between the 1194th A and the 1195th T of the sequence. The MPTP-delta
FT sequence was isolated from murine brain tissue and was cloned, for
FT expression, into the downstream region of a glutathione-S-transferase
FT sequence and expressed as a fusion protein in E. coli. MPTP-delta
```

CC proteins regulate differentiation and activation of cells. This
 CC sequence can be used in the elucidation of the molecular mechanism for
 CC information transmission in cells, regulation mechanisms in the nervous
 CC system or immune system, or in the mechanism of carcinogenesis.
 SQ Sequence 225 AA;
 R75202 Length: 225 February 11, 2000 15:48 Type: P Check: 824 ..
 1 MPTNCRVVP VRRPISLLT FFLCACETP PRFTRPVQ TGVSGVASF
 51 ICAATGDRP KIYNNKGGK VSNQREYIE FDDGSSYLK IOPLEPRDE
 101 AIVECVASNN VGEISVSTR TYLRDQDPR GFPTIDMPQ LKVERTTA
 151 TWLCAASNP DPEITWFKP LVPDTSNNG RIKQRESEI GALQEGSEE
 201 SDQKRECV TNSAGTRISA PANLY
 !!AA SEQUENCE 1.0
 ID R91274 standard; Protein; 62 AA.
 AC R91274;
 DT 18-JUN-1996 (first entry)
 DE Elmeria gametocyte antigen.
 KW Vaccine; coccidiosis; chicken; hen; fowl; immunisation;
 KW maternal immunity; gametocyte; antigen.
 OS Elmeria maxima.
 PN US5496550-A.
 PD 05-MAR-1996.
 PF 14-AUG-1986; 896611.
 PR 14-AUG-1986; US-896611.
 PR 12-FEB-1988; US-155245.
 PR 14-FEB-1989; US-310603.
 PR 16-JAN-1991; US-642219.
 PR 17-AUG-1993; US-108763.
 RA (CHIT-) CHITMALNER.
 RA Wehner D, Pugatsch T, Wallach M;
 PI WPI; 96-150679/15.
 DR N-PSDB; T13968.
 DT Protection of newborn chicks from Elmeria spp. infection - by
 PT immunisation of laying hen prior to the hen laying with an Elmeria spp.
 PT gametocyte antigenic protein
 PS Example 9; Fig 14a; 3pp; English.
 CC Clone PEM 36/3 (ATCC 68228) represents the 3' half of either a 2.0
 CC or 3.5 kb mRNA of Elmeria maxima gametocytes. It was obtd. by
 CC screening an E. maxima cDNA library using mouse polyclonal antibodies
 CC to the 82 kDa gametocyte antigen. There are 2 possible open reading
 CC frames (T13968 and T13969) coding for different proteins (R91274 and
 CC R91275, respectively). Cloning of Elmeria gametocyte antigen genes
 CC (see also R91271-73) will allow large scale production of the antigens,
 CC useful for immunising hens to protect newborn chicks against
 CC coccidiosis.
 SQ Sequence 62 AA;
 R91274 Length: 62 February 11, 2000 15:48 Type: P Check: 1456 ..
 1 LQHHQHITL QOOQHHPH PAAVLVAPR ADEKQFRAA AAPPIFAA
 51 SKETPLIDMM SK
 !!AA SEQUENCE 1.0
 ID R94989 standard; Peptide; 354 AA.
 AC R94989;
 DT 06-JUL-1996 (first entry)
 DE Nsk2 receptor intracellular domain (aa518-871).
 KW Nsk2; receptor fold/somite kinase; receptor tyrosine kinase; muscle;
 KW vector; tumour; muscular dystrophy; diagnosis; therapy.
 OS Mus musculus.
 FH Key
 FH region location/Qualifiers
 FT 1..59
 FT region /label= juxtamembrane_region
 FT 60..341
 FT region /label= Tyrosine-kinase_region
 FT 157..177
 FT region

FT region /label= Kinase_insert
 FT 342..354
 FT region /label= C-terminal_region
 PN WO9611664-A2.
 PD 25-APR-1996.
 PR 10-OCT-1995; U13490.
 PR 10-OCT-1994; GB-020389.
 PR 24-MAY-1995; GB-010574.
 RA (LUDW-) LUDWIG INSR CANCER RES.
 PI Reith A;
 PI WPI; 96-221725/22.
 DR Nucleic acid encoding muscle receptor tyrosine kinase(s) - and
 PT related vectors, transformed cells, polypeptide(s) and antibodies,
 PT useful for treating tumours, diagnosing muscular dystrophy, etc.
 PS Claim 8; Fig 2; 65pp; English.
 CC The intracellular domain (R94990-92) of mouse Nsk2 (R84087) can be
 CC (see also R94981-88 and R94990-92) of mouse Nsk2 (R84087) can be
 CC used to study the role of Nsk2 in growth and differentiation, in
 CC the development of Nsk (ant)agonists to identify other components
 CC of Nsk signalling pathways, and in the diagnosis and therapy of e.g.
 CC tumours and muscular dystrophy. Nsk2 is a novel member of a new
 CC family of muscle receptor tyrosine kinases that are expressed in
 CC terminally differentiated skeletal myotubes in vitro and in adult
 CC skeletal muscles in vivo.
 SQ Sequence 354 AA;
 R94989 Length: 354 February 11, 2000 15:48 Type: P Check: 5431 ..
 1 YCCRRRERK NKRREYAVT LTLPEELL DRLHPNMQ RMLPLNKL
 51 LSLFYPRNNI EYVDIEGA FGRVQARAP GLIPEFTM VAVMKKEEA
 101 SADNQDFOR EALMAEFDN PNIVKLGVC AVGRPMCLF EYMAVDINE
 151 FLRMSPHTV CSLSHDST RARVSSGPP PLSCAOLCI AQVAGNAY
 201 LSRKRVHND LATNCLYGE TMVVKIADFG LSRNYSADY YKADNDALP
 251 IRKMPESIF YNRRTESD WAGVYLWEI FSYGLQPYG MAHEVIYVY
 301 RDGNILACP NCPLELYLM RLCWSLPPAD RPSFSIHRI LQMRERAG
 351 TVGV
 !!AA SEQUENCE 1.0
 ID R94990 standard; Peptide; 282 AA.
 AC R94990;
 DT 06-JUL-1996 (first entry)
 DE Nsk2 receptor tyrosine-kinase region (aa577-858).
 KW Nsk2; neutral fold/somite kinase; receptor tyrosine kinase; muscle;
 KW vector; tumour; muscular dystrophy; diagnosis; therapy.
 OS Mus musculus.
 PN WO9611664-A2.
 PD 25-APR-1996.
 PR 10-OCT-1995; U13490.
 PR 10-OCT-1994; GB-020389.
 PR 24-MAY-1995; GB-010574.
 RA (LUDW-) LUDWIG INSR CANCER RES.
 PI Reith A;
 PI WPI; 96-221725/22.
 DR Nucleic acid encoding muscle receptor tyrosine kinase(s) - and
 PT related vectors, transformed cells, polypeptide(s) and antibodies,
 PT useful for treating tumours, diagnosing muscular dystrophy, etc.
 PS Claim 8; Fig 2; 65pp; English.
 CC The tyrosine kinase region (R94990) and other polypeptide portions
 CC (see also R94981-88 and R94991-92) of mouse Nsk2 (R84087) can be
 CC used to study the role of Nsk2 in growth and differentiation, in
 CC the development of Nsk (ant)agonists, to identify other components
 CC of Nsk signalling pathways, and in the diagnosis and therapy of e.g.
 CC tumours and muscular dystrophy. Nsk2 is a novel member of a new
 CC family of muscle receptor tyrosine kinases that are expressed in
 CC terminally differentiated skeletal myotubes in vitro and in adult
 CC skeletal muscles in vivo.

Mon Feb 14 08:07:18 2000

ags.cat

Page 42

SQ Sequence 282 AA;

R94990 Length: 282 February 11, 2000 15:48 Type: P Check: 5928

1 IEVYRDGEG AGRVFOQARA POLLEPEPPT MVAVKMLKEE ASADQADFO
51 REALMAEFD NPVIVKLGV CAVGKPMCLL FEYWAYGDILN EFLRMSPT
101 VCSISHDLS TRAVSSPGE PPLSCAEQLC IARQVAGNA YLSERKFVHR
151 DLAIRNLGV EIMVVKIADF GLSRNITSAD YKADNDAI PIRWAPPEI
201 FYRRTTESD VNAVGVLEWE IFSYGLOPY YGMAHEVIYY VRDGNILACP
251 ENCPLEIYNL MRLCWSKLPA DRPSFCISHR IL

11AA_SEQUENCE 1.0
ID R89896 standard; Protein: 423 AA.
AC R89896;

DE 14-JUL-1996 (first entry)
DE Rat kynurenine aminotransferase (KAT) clone.
KW Kynurenine aminotransferase; KAT; kynurenic acid; KYNA; kynurenine;
OS Rattus rattus.
PN WO9601893-A1.
PD 25-JAN-1996.
PF 23-JUN-1995; U07855.
PR 07-JUL-1994; US-271667.
PA (PHAA-) PHARMACIA SPA.
PA (UTMA-) UNIV MARYLAND BALTIMORE.
PI Benatti L, Breton J, Mosca M, Okuno E, Schwarcz R;
PI Speciale C;
DR N-PSDB: T11742.
DR WPI: 96-097623/10.
PT Isolated DNA encoding mammalian kynurenine amino:transferase (KAT) -
PT useful in gene therapy applications and for identifying KAT in brain
PT tissue
PS Claim 16: Figure 2: 51pp: English.
CC Sequences encoding Kynurenine aminotransferase (KAT) can be inserted
CC into vectors and subsequently cells and hence can be used for gene
CC therapy. The vector and host cells can be used for cerebral
CC implantation to where KAT can directly catalyse the production of
CC kynurenic acid (KYNA) from kynurenine (KYN). It is thought KYNA acts
CC as a negative endogenous modulator of cerebral glutamatergic
CC function. KYNA concentrations and the activity of KAT show an
CC increase with age. KAT inhibitors, by providing an increase of the
CC glutamatergic tone at the NMDA receptor, could be useful in
CC situations where NMDA receptor function is insufficient and/or KAT
CC activity and KYNA levels are abnormally enhanced. Hence they could
CC be particularly useful in the treatment of the pathological
CC consequences associated with the aging processes in the brain.
CC Three KAT clones are described in T11560, T11742-43.
SQ Sequence 423 AA;

R89896 Length: 423 February 11, 2000 15:48 Type: P Check: 6122

1 MTRKLRARL DSDIDNLWE FGKLTKEYDV VNLGGFPDF SPPEFAQAF
51 QOATSGNFMML NOYTRAFGYP PLTNVLASF GKLIGQEMDP LTNVLYVGA
101 YGALTFQFA LVDEGDEVII MEPAFDCEP MTMMAGGCY FVITKESPPAP
151 KGRIGASNDW QLDPAELASK FTPTKTYLV NTPNNPLGK FSRRELEVA
201 NLCOQHDVVC ISDEVYQWLV YDGHQVSTA SLPGMDRTL TISGAKSFS
251 ATGWKVGWVA GPDNIMKLR TVQNSIFHC PTOQAAYAAQ CFEREQOHFG
301 OPSSTYLOLP QAMELNBDHM IRSLOVGLK LWISGSYFL IADISPFKS
351 MPDLGAEDE PYDRRFKMM IKMGLVGLIP VSTFSPRHQ KDFDHYIRFC
401 FKDKATILQA MDERLRKWE LOP

11AA_SEQUENCE 1.0
ID R89897 standard; Protein: 437 AA.
AC R89897;

DE 14-JUL-1996 (first entry)
DE Rat kynurenine aminotransferase (KAT) clone.
KW Kynurenine aminotransferase; KAT; kynurenic acid; KYNA; kynurenine;
OS Rattus rattus.
PN WO9601893-A1.
PD 25-JAN-1996.
PF 23-JUN-1995; U07855.
PR 07-JUL-1994; US-271667.
PA (PHAA-) PHARMACIA SPA.
PA (UTMA-) UNIV MARYLAND BALTIMORE.
PI Benatti L, Breton J, Mosca M, Okuno E, Schwarcz R;
PI Speciale C;
DR N-PSDB: T11742.
DR WPI: 96-097623/10.
PT Isolated DNA encoding mammalian kynurenine amino:transferase (KAT) -
PT useful in gene therapy applications and for identifying KAT in brain
PT tissue
PS Claim 16: Figure 3: 51pp: English.
CC Sequences encoding Kynurenine aminotransferase (KAT) can be inserted
CC into vectors and subsequently cells and hence can be used for gene
CC therapy. The vector and host cells can be used for cerebral
CC implantation to where KAT can directly catalyse the production of
CC kynurenic acid (KYNA) from kynurenine (KYN). It is thought KYNA acts
CC as a negative endogenous modulator of cerebral glutamatergic
CC function. KYNA concentrations and the activity of KAT show an
CC increase with age. KAT inhibitors, by providing an increase of the
CC glutamatergic tone at the NMDA receptor, could be useful in
CC situations where NMDA receptor function is insufficient and/or KAT
CC activity and KYNA levels are abnormally enhanced. Hence they could
CC be particularly useful in the treatment of the pathological
CC consequences associated with the aging processes in the brain.
CC Three KAT clones are described in T11560, T11742-43.
SQ Sequence 437 AA;

R89897 Length: 437 February 11, 2000 15:48 Type: P Check: 206

1 MNSTWCLFF KRLNTRKRLQ ARRLLGIDON LWEFGKLT EYDVNLGOG
51 FPDSPEDRA TOAFQOATS NEMLNQYTRA EGYPLTNVL ASFGKLLQG
101 EMDPLTNVL TVGAYGALT AFOALVDEGD EYIMEPAFD CYEPMTMAG
151 GCPEVYTLRP SPAPKGLGA SMDQIDPAE LASKTPTKX VLVLNTPNP
201 LGRVFSRML ELVANTCOOH DVVCISDEVY QWLYVDGHH VSIASLPGW
251 DRTLIGSAG KFSFATGWKY GWVWGPDNIM KHLRTVHNS IIFCPTQAOA
301 AVAOCFEREO QHFGOPSSYF LQLPQAMELN RDHMRISLOS VGLKLWISG
351 SYFLADIAD FSKRMPDLPG AEDEPYDRRF AKMWIKMGL VGIPVSTFS
401 RPHOKEDHY IRFCVAKDA TLQANDERL KMKELQ

11AA_SEQUENCE 1.0
ID R89898 standard; Protein: 457 AA.
AC R89898;

DE 14-JUL-1996 (first entry)
DE Rat kynurenine aminotransferase (KAT) clone.
KW Kynurenine aminotransferase; KAT; kynurenic acid; KYNA; kynurenine;
OS Rattus rattus.
PN WO9601893-A1.
PD 25-JAN-1996.
PF 23-JUN-1995; U07855.
PR 07-JUL-1994; US-271667.
PA (PHAA-) PHARMACIA SPA.
PA (UTMA-) UNIV MARYLAND BALTIMORE.

Mon Feb 14 08:07:18 2000

ags.cat

Page 44

SO Sequence 342 AA;

R95054 Length: 342 February 11, 2000 15:48 Type: P Check: 9286

```
1 DYKDDDDKLG TGVVSHFNDG PDSHTQFCFH GTCRFVQED KPAVCYCHSGY
51 VGARCEHADL LASLEHHHHH HLEGGSLAL TAHQACHLPL EFTFRHQRPR
101 GWEOLEOCGY PVORLVALYI AARLSMNOVD QVIRNALASP GSGGDGGEAI
151 RQEPQEARLA LTLAAESER FVROGTGNDE AGANADEKL LSTIQACDI
201 CRKKKCKSK EKPRCAKCKR NMECRSPK TKRSPILRAH LVEFSRLER
251 LEQFLILFP REDLDMLRM DSLQDIKALL TGLFQDVNVA KQAVTDRLAS
301 VEIDMPLTLR QHRISATSSS EESSNKGQRQ LTVSSSDYMD EL
```

!!AA_SEQUENCE 1.0
ID R95055 standard; Protein: 421 AA.

AC R95055; (first entry)
DE IL-2-DETA-DGAL4 multidomain protein.
KW Nucleic acid transfer system; gene transfer; gene therapy;
KW cell targeting; multidomain protein; vector; cancer;
KW exotoxin A; DETA; signal peptide; GAL4; interleukin-2;
IL-2.

OS Chimeric synthetic;
OS Chimeric Homo sapiens;
OS Chimeric Pseudomonas aeruginosa;
OS Chimeric Saccharomyces cerevisiae;
FH Key Location/Qualifiers

```
1..8 /label= FLAG-epitope
FT peptide 9..17 /label= Spacer
FT 18..150 /label= IL-2
FT /note= "amino acids 1-113 of human IL-2"
FT peptide 151
FT /label= Spacer
FT 152..266 /label= ETA
FT /note= "amino acids 252-366 of ETA"
FT domain 267..413 /label= Spacer
FT 268..413 /label= GAL4
FT /note= "amino acids 2-147 of yeast GAL4"
FT peptide 414..421 /label= Spacer
FT /note= "endoplasmic reticulum retention signal"
```

MO9613599-A1.
PD 03-MAY-1996.
PR 31-OCT-1995; E04270.

PA 01-NOV-1994; EP-810627.
PI (WELLS/) WELLS W;
PI Fenninaya J; Wells W;
DR N-PSDB; T239411.

PI Nucleic acid transfer system for gene therapy, e.g. against cancer
PI - includes toxin translocation domain to target nucleic acid to
PI specific cell

PS Claim 7; Page 67-69; 106PP; English.
CC A multidomain protein (R95055) has a FLAG epitope, a portion
CC of human interleukin-2 that acts as a ligand domain, a
CC non-cytotoxic portion of Pseudomonas aeruginosa exotoxin A acting
CC as a translocation domain and the DNA binding domain of yeast GAL4.
CC It is the product of a fusion gene (T29411) and can be expressed
CC in E. coli (resulting in removal of an ompA signal peptide). It is
CC used with an effector nucleic acid that comprises e.g. a gene to be
CC delivered to a cell and a cognate structure for the GAL4 DNA binding
CC domain. This provides a novel means of nucleic acid transfer.

CC suitable for gene therapy.
SQ Sequence 421 AA;

R95055 Length: 421 February 11, 2000 15:48 Type: P Check: 7433

```
1 DYKDDDDKHL HHHHKKLAPT SSSKTKTQQL LEHLLDDLM ILGNNYKN
51 FKLRMLTER FYMPKATKL KHLQCLEEEL KPLEEVLVA QSNFHLRPR
101 DLISNNIVYI LEKGETTF MCEYADETN IVEFLNRIIT FQSIISTLT
151 LEGGSIALTL AHQACHLEPLE TFRHQRPRG WEOLEOCGYR VORLVALYIA
201 ARLSMNOVDQ VIRNALASP GSGGDGGEAIR EQPQARLAL TLAAESERF
251 VROGTGNDEA GAANADEKL SSIEQACDIC RLKRLCKSKE KPRCAKCLKN
301 NMECRSPKTKRSPILRAHL TEVESRLERL EQFLILIFPR EDLDMLKMD
351 SLQDIKALLT GLFQDVNVA KQAVTDRLASV ETDMPILTRQ HRISATSSSE
401 EESSNKGQRQL TVSSSDYKDE L
```

!!AA_SEQUENCE 1.0
ID R91225 standard; Protein: 328 AA.
AC R91225; (first entry)

DE Human placenta G-protein coupled receptor protein.
KW G-protein coupled receptor protein; G-PCR; agonist; antagonist;
KW cystic fibrosis; incontinence; diabetes; diagnosis; therapy.
OS Homo sapiens;
MO9605303-A1.
PD 22-FEB-1996.
PF 10-AUG-1995; J01589.
PR 11-AUG-1994; JP-189272.
PR 11-AUG-1994; JP-189273.
PR 11-AUG-1994; JP-189274.
PR 30-SEP-1994; JP-236356.
PR 02-SEP-1994; JP-236357.
PR 02-NOV-1994; JP-270017.
PR 28-DEC-1994; JP-326611.
PR 10-UNK-1995; JP-007117.
PR 16-MAR-1995; JP-057186.
PR 15-APR-1995; JP-093989.

PA (TAKA) TAKEDA CHEM IND LTD.
PI Fujii R; Fukusumi S; Hinuma S; Hosoya M; Ohgi K;
PI Ohtaki T;
DR WPI; 96-139698/14.
DR N-PSDB; T18368.
PT G-protein coupled receptor protein DNA and protein - also methods
PT for isolating (ant)agonists for treatment of cystic fibrosis,
PT incontinence and diabetes

PS Claim 6; Page 267-68; 360PP; English.
CC A novel human placenta-derived G-protein coupled receptor protein
CC (G-PCR) (R91225) was identified as the product of cDNA clone
CC pHAH2-17 (T18368). The protein can be obtained by expression of the
CC cDNA clone in transformed host cells. It was classified as a
CC purinoceptor. G-PCRs (see also R91217-24 and R91227-33) can be
CC used to screen agonists and antagonists that modulate G-PCR
CC activity, to raise antibodies and to develop assay systems.

SQ Sequence 328 AA;
R91225 Length: 328 February 11, 2000 15:48 Type: P Check: 7811

```
1 MEMONGGQA LGPPTTCY RENEKOLLP PYSAVLANG LPLNICVITQ
51 ICTSRRLTR TAYTINLAL ADDLYACSLP LLIYNAQGD HMPGDFACR
101 LVRLFLANL HGSILFLCI SFORTIGICH PLAPWKRKG RRAWLVCYT
151 VWLAVTQCL PPAIPAATGI QRNRVWCYDL SPALATHYM PYGALVIG
201 FLDPFALLA CYCLACRLC RODGPAEPVA QRRRGAARM AVVYAAFAI
```


251 SELEPHITKT AYLVAGSTPG VPCTVLEAFV AAYKGRPFA SANSYLDPLT
 301 FYTOKKRR RHPELLOKLT AKMORGR

11AA-SEQUENCE 1.0
 ID R99695 standard; Peptide: 25 AA.
 AC R99695;
 DT 25-SEP-1996 (first entry)
 DE VP6 binding peptide B; encapsulation; target; agent; drug delivery;
 KW VP6; binding peptide; encapsulation; target; agent; drug delivery;
 KM virus protein 6; rotavirus; specific.
 OS Synthetic.
 PN PT-100090-A.
 PD 30-JUL-1993.
 PE 23-DEC-1992; 100090.
 PR 04-FEB-1991; US-650054.
 PR 29-JAN-1992; US-825522.
 PR 06-SEP-1994; US-301267.
 PR (UYSA-) UNIV SASKATCHEWAN.
 PA Redmond MJ, Santos M, Campos M;
 PI WPI: 93-270390/34.
 DR Encapsulated bioactive protein compns. - contg. protein in
 PT rotavirus VP6 protein spheres
 PS Dislosure; Column 9; 12pp; English.
 CC VP6 (is a 45 kD protein found in the inner capsid of rotaviruses. VP6
 CC rapidly targets, attaches to and is engulfed by monocytes and
 CC macrophages. Substances to be delivered to target cells can be
 CC encapsulated in VP6 spheres to effect an increase in the concentration
 CC of the agent delivered to the cell. Additionally, other targeting
 CC agents can be attached to the VP6 spheres that will direct the spheres
 CC to other tissues and cell types so that biologically active agents
 CC encapsulated within the spheres can act locally. An advantage of the VP6
 CC protein is that this system facilitates the attachment of molecules with
 CC minimal manipulation, through protein-protein interactions. A synthetic
 CC peptide corresponding to the targeting agent of interest can be
 CC chemically synthesised in such a way that it also contains an amino acid
 CC sequence (binding peptide) necessary to link it to VP6. Attachment of the
 CC molecule to the VP6 carrier is then simply achieved by mixing the 2
 CC substances without additional manipulation. The present sequence is a
 CC peptide that binds to VP6. It occurs naturally as a portion of VP4 of
 CC rotaviruses and is sensitive to trypsin. Cleavage of the peptide by
 CC trypsin prevents it from binding to VP6. Patent No. US5503833 was used
 CC to compile this file.
 SQ Sequence 25 AA;

R99695 Length: 25 February 11, 2000 15:48 Type: P Check: 4995 ..

1 CNIPASIVS RNIVYTRAQP NODIA

11AA-SEQUENCE 1.0
 ID R99501 standard; Protein: 244 AA.
 AC R99501;
 DT 14-SEP-1996 (first entry)
 DE Nitrilase regulatory factor 244-amino acid subunit.
 KW Nitrilase regulatory factor; gene activator; promoter; nitrile.
 OS Rhodococcus erythropolis strain SK92 (FERM BP-3324).
 PN EP-719862-A2.
 PD 03-JUL-1996.
 PR 27-DEC-1995; 309454.
 PR 28-DEC-1994; JP-337652.
 PR (NITR) NITRO CHEM IND CO LTD.
 PA Mizumura Y, Yu F;
 PI WPI: 96-302346/31.
 DR N-PSDB: T32242.
 DR Rhodococcus erythropolis two component nitrilase gene activator
 PT enhances activation of nitrilase gene promoter in presence of
 PT nitrile
 PT Claim 1, Page 9-10; 23pp; English.
 PS The nitrilase regulatory factor of Rhodococcus erythropolis SK92
 CC has 2 subunits, of 244 amino acids (R99501) and 534 amino acids
 CC (R95502). These are the products of the 2 open reading frames
 CC (see also T32242 and T32243) of a gene (T32241) located upstream

of the nitrilase gene in SK92 chromosomal DNA. Introduction of a
 CC plasmid carrying the regulatory gene plus the nitrilase gene and its
 CC promoter allows efficient prodn. of nitrilase by a Rhodococcus sp. of
 CC transformant. The nitrilase is useful for the commercial prodn. of
 CC organic acids and amides from nitriles.
 SQ Sequence 244 AA;

R99501 Length: 244 February 11, 2000 15:48 Type: P Check: 2030 ..

1 MAGADYHAG GTRNARILV VDDEKHVRM VTMOLESENF DVAAADGDA

51 ALRQVTESEP DLMVLDLSP GKGGEVLAT VRRDALPIV VLIARDETE

101 RIVALDLDGD DYIKPFSR ELAARAVL RRTAEPPE AAVORFGDLE

151 IDTAREVRL HGPLEFTTK EFDLALYMA SPQGVFERR LLEVMRSSP

201 DWQDAVTE HVHRIRKIE EDETKPTIQ TVRAGYRFD GERA

11AA-SEQUENCE 1.0
 ID R99576 standard; Protein: 83 AA.
 AC R99576;
 DT 30-SEP-1996 (first entry)
 DE Wasp venom BrhX-1 subunit (b) partial sequence.
 KW Wasp; venom; neurotoxin; insecticide; biological control agent;
 KM Lepidoptera; Insect.
 OS Braccon hebetor.
 PN W09616171-A1.
 PD 30-MAY-1996.
 PE 21-NOV-1995; G02720.
 PR 22-NOV-1994; GB-023540.
 PR 19-JAN-1995; GB-001074.
 PR 29-JUN-1995; GB-013293.
 PR (CSTR) COMMONWEALTH SCI & IND RES ORG.
 PA (ZENE) ZENCA LTD.
 PI Baile VJ, Christian PD, Duncan RE, Windass JD;
 DR WPI: 96-268607/27.
 DR N-PSDB: T32444.
 DT Braccon hebetor toxins and DNA encoding them - useful in biological
 PT control agents to combat insect pests
 PS Example 6, Fig 11; 83pp; English.
 CC Partial cDNA clone pbhX-1(b)1 (T32444) codes for a portion
 CC (R99576) of subunit (b) of the Braccon hebetor wasp neurotoxin
 CC BrhX-1. It was obtained by PCR amplification of wasp cDNA in
 CC a phage lambda gIII library and characterisation of a plaque-purified
 CC phage clone. Primers BH(b) and BH(b)D (T32445-46) based on the clone
 CC were used to re-screen the library leading to the isolation of clone
 CC BrhX-1(b) (T32429) that coded for the toxin (b) subunit (R99577).
 CC This can be utilised in breeding of biological control agents used to
 CC combat insect pests.
 SQ Sequence 83 AA;

R99576 Length: 83 February 11, 2000 15:48 Type: P Check: 2766 ..

1 FFLITSIIC KILLVLLSW TSMVSTLFT DRKMGGRADK TFGPERSLGG

51 GVGDCCRSHD SCGRMKRPE TGDVYTKNGF SNM

11AA-SEQUENCE 1.0
 ID W00624 standard; Protein: 358 AA.
 AC W00624;
 DT 23-OCT-1996 (first entry)
 DE Saci endonuclease.
 KW Saci; endonuclease; Streptomyces achromogenes; Saci; methylase; endonuclease.
 OS Streptomyces achromogenes.
 PN US5532153-A.
 PD 02-JUL-1996.
 PR 23-MAR-1995; 409199.
 PR (NEME) NEW ENGLAND BIOLABS INC.
 PA Xiao J, Xu S;
 PI WPI: 96-321120/32.
 DR N-PSDB: T35805.

PT DNA encoding SacI restriction endonuclease and methylase - useful
 for prod. of the enzyme.
 PS Claim 3: Column 15-18; 16pp; English.
 CC This sequence represents the SacI endonuclease from *Streptomyces*
 CC achromogenes. W00623 represents the SacI methylase. The DNA encoding
 CC these sequences were isolated from isolated from genomic DNA of
 CC *S. achromogenes* and amplified using the primers shown in 135806-135811.
 CC The amplified sequences were then cloned into the plasmid pG339 which
 CC was used to modify an *E. coli* host cell. The host cell can then be used
 CC to produce SacI.
 SQ Sequence 358 AA.

W00624 Length: 358 February 11, 2000 15:48 Type: P Check: 8391 ..

1 MGTTRKSTRA BOVRKAYEA AASDVLELD WIFLATSIRE VDAFRTYTA
 51 LVTALIRAC DDRVDPRSIK EKYDDRAFSL RILCHGVAVP MSVELGPDLC
 101 ATGREPINNO PFRRDYDSE IYRVOIKRBP YIDRVSSALA RVDEEDVSTE
 151 ESFRALVAVL AVCSIVANKR QRYAVGSAIV EASLIAETQS FVVGSDVPR
 201 KLOACVAAGL DMVSEYVSR RINDPDRDP GDVOYILDG PLLTVEGRK
 251 SVSWEGLBOF VSSATYAGFR RVALWYDAS HVSLMSADL TSLERKRYEC
 301 IVKNESVYS FLRDVEWSP RDVHSILSAF PEANYRMIE IYRREPELDR
 351 WAETPFET
 !!AA_SEQUENCE 1.0
 ID R94579 standard; Protein: 488 AA.
 AC R94579;
 DT 07-NOV-1996 (first entry)
 DE Chlamydia pneumoniae polypeptide antigen (polypeptide A).
 KW Polypeptide antigen; polypeptide A; strain Yk41; plasmid; probe;
 KM PCPN533alpha; primer; assay; detection; antibody; diagnosis;
 OS Chlamydia pneumoniae.
 FT Key location/Qualifiers
 FT 1..259
 FT peptide
 FT W09609320-A1.
 PD 28-MAR-1996.
 PF 20-SEP-1995; J01896.
 PR 28-APR-1995; JP-224711.
 PR 28-APR-1995; JP-106011.
 PR 28-APR-1995; JP-106006.
 PR 28-APR-1995; JP-106010.
 PR 28-APR-1995; JP-106009.
 PA (HITB) HITACHI CHEM CO LTD.
 PI Izutsu H, Matsumoto A, Obara K;
 DR WPI: 96-188399/19.
 N-PSDB: T14612.
 PS Recombinant Chlamydia pneumoniae antigen and antibodies to it
 PS used for detection and assay of *C. pneumoniae* e.g. in clinical
 PS diagnosis
 PS Claim 1: Pages 60-64; 128pp; Japanese.
 CC The present sequence is the *C. pneumoniae* polypeptide antigen,
 CC polypeptide A. *C. pneumoniae* strain Yk41 was cultured and genomic
 CC DNA extracted to prep. a lambda gtl1 DNA library. The library was
 CC then screened with an anti-Yk41 monoclonal antibody (MAb), which
 CC was prepd. by fusing spleen cells from a mouse infected with Yk41
 CC with myeloma P3/NS1/1-Ag4-1 to produce a MAb expressing hybridoma.
 CC The DNA of the MAb was then fused with the expression vector pAD431 to
 CC give PCPN533alpha. The plasmid was used to transform an *E. coli*
 CC host, which was then fused with the antigenic polypeptide A
 CC polypeptide A. Polypeptide A and primers and probes derived from
 CC its DNA can be used in assays for the detection of polypeptide A
 CC antibodies and DNA, respectively, useful in the diagnosis of
 CC *C. pneumoniae* infection.
 SQ Sequence 488 AA.

R94579 Length: 488 February 11, 2000 15:48 Type: P Check: 2400 ..

1 MTSISSSGPD NOKNMSOVL TSTPOGVPOO DKLSGNETKQ IOOTROGKNT
 51 EHSDDATLNG ASGRDKTSST TKETAPPOG VAAKSESSES OKAGADTGS
 101 GAATTAASNT ATKIAMQTSI EEAASKMEST LESLOSLSAA OKKEVEAVVY
 151 AALSCKSSGS AKLETPELPR PGVTPRSEVI EIGLALAKAI OLGENTATSA
 201 LSNVASTOAO ADOTNKLGL EKOAIKIDKER EYEOEKMAE OKSKLEGTM
 251 DTVNTMIAN SVAITIISIV AAFICGAGL AGLAABAARV AAAAGGAGA
 301 AAATTVATQI TVQAVQAVK QAVITAVRQA ITAIAKAARV SGIRAFITL
 351 VKAIAKAISK GISKVPAKGT QMIANKPEKL SKYISLSISK WYVGVGVVY
 401 AAPALGKIM OMOLSEMOON VAOFQKEVK LQAADIMISM FTORWOQASK
 451 IASKOTGESN EMTOKATKLG AOLKAYPAI SGAIAGAA
 !!AA_SEQUENCE 1.0
 ID R94580 standard; Protein: 271 AA.
 AC R94580;
 DT 07-NOV-1996 (first entry)
 DE *C. pneumoniae* polypeptide antigen (polypeptide A) variant.
 KW Polypeptide antigen; polypeptide A; strain Yk41; plasmid; probe;
 KM PCPN533alpha; primer; assay; detection; antibody; diagnosis;
 OS Chlamydia pneumoniae.
 FT Key location/Qualifiers
 FT 1..271
 FT peptide
 FT W09609320-A1.
 PD 28-MAR-1996.
 PF 20-SEP-1995; J01896.
 PR 28-APR-1995; JP-224711.
 PR 28-APR-1995; JP-106011.
 PR 28-APR-1995; JP-106006.
 PR 28-APR-1995; JP-106010.
 PR 28-APR-1995; JP-106009.
 PA (HITB) HITACHI CHEM CO LTD.
 PI Izutsu H, Matsumoto A, Obara K;
 DR WPI: 96-188399/19.
 N-PSDB: T14613.
 PS Recombinant Chlamydia pneumoniae antigen and antibodies to it
 PS used for detection and assay of *C. pneumoniae* e.g. in clinical
 PS diagnosis
 PS Claim 2: Pages 64-66; 128pp; Japanese.
 CC The present sequence is a variant of the *C. pneumoniae*
 CC polypeptide antigen, polypeptide A. *C. pneumoniae* strain Yk41 was
 CC cultured and genomic DNA extracted to prep. a lambda gtl1 DNA
 CC library. The library was then screened with an anti-Yk41
 CC monoclonal antibody (MAb), which was prepd. by fusing spleen cells
 CC from a mouse infected with Yk41 with myeloma P3/NS1/1-Ag4-1 to
 CC produce a MAb expressing hybridoma. The DNA of the MAb was then fused
 CC with the expression vector pAD431 to give PCPN533alpha. The
 CC plasmid was used to transform an *E. coli* host, which was then fused
 CC to give the antigenic polypeptide A. Polypeptide A and primers and probes
 CC derived from its DNA can be used in assays for the detection of polypeptide A
 CC antibodies and DNA, respectively, useful in the diagnosis of *C. pneumoniae*
 CC infection.
 SQ Sequence 271 AA.

R94580 Length: 271 February 11, 2000 15:48 Type: P Check: 3773 ..

1 MTSISSSGPD NOKNMSOVL TSTPOGVPOO DKLSGNETKQ IOOTROGKNT
 51 EHSDDATLNG ASGRDKTSST TKETAPPOG VAAKSESSES OKAGADTGS
 101 GAATTAASNT ATKIAMQTSI EEAASKMEST LESLOSLSAA OKKEVEAVVY
 151 AALSCKSSGS AKLETPELPR PGVTPRSEVI EIGLALAKAI OTIGENTATSA

201 LSNVASTQAO ADQTNKLGLE KOAIKIDKER EYEOEKMAE OKSKDEGTM
 251 DTVNTVMIAK GFELWGPPLI N

!!AA SEQUENCE 1.0
 ID R94586 standard; Protein: 259 AA.
 AC R94586
 DE C. pneumoniae polypeptide antigen (polypeptide A) clone 53-35.
 KW Polypeptide antigen; polypeptide A; strain YK41; plasmid; probe;
 NM pN333alpha; primer; assay; detection; antibody; diagnosis;
 OS Chlamydia pneumoniae.
 PS W09609320-A1.
 PD 28-MAR-1996.
 PF 20-SEP-1995; J01896.
 PR 28-SEP-1994; JP-224711.
 PR 28-APR-1995; JP-106011.
 PR 28-APR-1995; JP-106006.
 PR 28-APR-1995; JP-106010.
 PR 28-APR-1995; JP-106008.
 PR 28-APR-1995; JP-106009.
 PA (HITB) HITACHI CHEM CO LTD.
 PI Izutsu H., Matsumoto A., Obara K;
 DR WPI: 96-188399/19.
 DR N-PSDB: T14622.
 PT Recombinant Chlamydia pneumoniae antigen and antibodies to it
 used for detection and assay of C. pneumoniae e.g. in clinical
 diagnosis

PS Example 1: Pages 75-77; 128pp; Japanese.
 CC The present sequence is the C. pneumoniae polypeptide antigen
 CC polypeptide A clone, 53-35. C. pneumoniae strain YK41 was
 CC cultured and genomic DNA extracted to prep. a lambda gtl1 DNA
 CC library. The library was then screened with an anti-YK41
 CC monoclonal antibody (mAb), which was prep. by fusing spleen cells
 CC from a mouse infected with YK41 with myeloma P3/NS1-1-Ag4-1 to
 CC produce a mAb expressing hybridoma. The DNA obd. was then fused
 CC with the expression vector pADA431 to give pCPN533alpha. The
 CC plasmid was used to transform an E. coli host which was cultured
 CC to give the antigenic polypeptide, polypeptide A. Polypeptide A
 CC and primers and probes derived from its DNA can be used in assays
 CC for the detection of polypeptide A antibodies and DNA,
 CC respectively, useful in the diagnosis of C. pneumoniae infection.
 CC Sequence 259 AA;

R94586 Length: 259 February 11, 2000 15:48 Type: P Check: 9728

1 MSISSSGPD NOKNMSOVL ISTPQGVPOQ DKLSGNETKO IQOTFROGANT

51 EMESDATING ASGRDXTSST TKTETAPQOG VAAKRESSSES OKAGADTGV

101 GAAATTSANT ATRIAMQTSI EASKSMEST LESIQSLSMA OKREVEAVV

151 AALSGKSSGS AKLEPPELPK PGVTPRSEVI EIGALAKAI OTLGATKSA

201 LSNVASTQAO ADQTNKLGLE KOAIKIDKER EYEOEKMAE OKSKDEGTM

251 DTVNTVMIA

!!AA SEQUENCE 1.0
 ID R99316 standard; Protein: 458 AA.
 AC R99316
 DE 13-NOV-1996 (first entry)
 DE Human SH-PTP1 variant derived from erythroleukemia cell line.
 KW PTP; protein tyrosine phosphatase; SH2; Src homology region 2;
 KW chromosome 12p; abnormality; mutation; detection; probe; neoplasia;
 KW cancer; leukemia; diagnosis; megakaryocyte regulation.
 OS Homo sapiens.
 PS U5536634-1.
 PD 16-JUL-1986.
 PD 26-JUN-1991; 721112.
 PR 26-JUN-1991; US-721112.

31-JAN-1992; US-829141.
 PR 01-DEC-1992; US-983926.
 PR 28-FEB-1994; US-202389.
 PA (BETH-) BETH ISRAEL HOSPITAL ASSOC.
 PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
 PI Freeman RM, Neel BG, Plutzky J, Rosenberg RD;
 DR WPI: 96-341506/34.
 DR N-PSDB: T35310.
 PT Detecting 12p chromosomal abnormality associated with neoplastic
 PT disease - using SH-PTP1 protein tyrosine phosphatase gene specific
 PT probe

PS Example 1: Column 49-52; 63pp; English.
 CC R99316 is a variant SH-PTP1 (protein tyrosine phosphatase-1, with
 CC two SH2 domains) obtained from a human erythroleukemia cell line.
 CC The variant cDNA sequence (see T35310) lacks a sequence of two
 CC nucleotides, AG corresponding to positions 1868 and 1869 of the
 CC wild-type sequence (see T35306). A fragment complementary to the
 CC sequence from nucleotides 537-653 is used as a probe to detect a
 CC chromosome 12p13 abnormality associated with neoplastic disease. In
 CC partic. acute lymphoblastic leukemia. The probe hybridises to a part
 CC of the region coding for the two tandem SH2 domains (see R99312).
 CC If the probe will not hybridise DNA of chromosome 12p13 from a patient
 CC sample it is indicative of an abnormality, normally associated with
 CC neoplasia. Alternatively the wild-type SH-PTP1 or SH-PTP2 gene or
 CC protein may be used for comparison to sequenced PTP genes taken from a
 CC patient, where differences indicate an abnormality. The activity of
 CC SH-PTP1 may also be purposefully altered by mutation to effect a change
 CC in megakaryocyte function and hence platelet production.
 CC Sequence 458 AA;

R99316 Length: 458 February 11, 2000 15:48 Type: P Check: 4323

1 MYRWTHRDLG GLDAETLLNG RGVHSGFLAR PSRNQDGF LSVRGDOVT

51 HIRIINSGDF YDIYGEKEFA TLELVEYTT QOQGVWDNRD GTIHLKPYL

101 NCSDFTERMR YGHMSGGQA ETILOAKGP WTELVRESLS OGDEVLAVL

151 SQPRAPGDS PLRVTHIKVM CEGRTYVGG LETPDSITLD VEHFKTIGE

201 EASGAFVLR QPYATRVNA ADIENRVLEL NKQSESDTA KAGFEFEFS

251 LKQEVKMLH QRELGQRPEN KGNRYKNIL PFDSRVITLQ GDSNIPGSD

301 YINAVYIKNO LIGPDENAKT YIASQCCELA TVNDFQMMAM QENSRYIVMT

351 TREVEKGRNK CEYNTITYP AMKNNAKAS RISKKHEDV YENLHKNNR

401 ESESEANVSR QGEQGFQPE EVSGAVLRMP CLSPDPEAF ROGQTHNLNL

451 GVPHSFVI

!!AA SEQUENCE 1.0
 ID R83015 standard; Protein: 371 AA.
 AC R83015
 DE 27-NOV-1996 (first entry)
 DE Human thyroid transcription factor-1.
 KW Thyroid transcription factor; TTF-1; human adenocarcinoma cell line;
 KW H441; rat; mouse; pulmonary adenocarcinoma; H820; small cell carcinoma;
 KW H345; tracheal-bronchial epithelial cell lines; respiratory epithelium;
 KW fetal lung; gestation; pro-SP-C; respiratory epithelial cell;
 KW nonciliated bronchiolar cell; immature lung; alveolar; Type II;
 KW epithelial cell; nonciliated; Type I; adult.
 OS Homo sapiens.
 PS W09531729-A1.
 PD 23-NOV-1995.
 PD 17-MAY-1995; U06244.
 PR 18-MAY-1994; US-245356.
 PA (CHIL-) CHILDRENS HOSPITAL MEDICAL CENT.
 PI Bohinski RJ, Whitsett JA;
 DR WPI: 96-011078/01.
 DR N-PSDB: T05810.

PT Oligo- and polynucleotide(s) that bind to lung cell nuclear proteins

PT - useful for cancer diagnosis and therapy.
 PS Claim 50; Fig 39; 157pp; English.
 CC This sequence represents the thyroid transcription factor, TTF-1. The
 CC TTF-1 DNA sequence was isolated from the human adenocarcinoma cell line
 CC H441. The TTF-1 locus is contained within a 4.6 kb BamH1 fragment and
 CC consists of two exons and one intron. This predicted amino acid
 CC sequence of human TTF-1 shows close identity with the amino acid
 CC sequence predicted by the rat TTF-1 cDNA sequence and 92.4% identity
 CC with the rat TTF-1 cDNA. The intron is approx. 1 kb in length and
 CC is flanked by consensus splice donor-acceptor sites that fit splice
 CC acceptor-donor rules. The mRNA produced is 2.3 kb as detected by
 CC Northern blot analysis of mRNA derived from mouse and human
 CC adenocarcinoma cells. TTF-1 mRNA was detected in human pulmonary
 CC not detected in 9/HTRCO- or BEAS-2B tracheal-bronchial epithelial cell
 CC lines, A549, Hela or 3T3 cells, demonstrating the cell selectivity of
 CC epithelium in human fetal lungs as early as 11-12 weeks of gestation.
 CC TTF-1 expression. TTF-1 has been detected in the developing
 CC respiratory epithelial cells in the developing lung, including
 CC airways similar to the for pro-SP-C. TTF-1 was detected in subsets of
 CC nonciliated bronchiolar cells and rarely in nonciliated bronchial respiratory
 CC epithelial cells in the mature lung. At the time of birth TTF-1 was
 CC detected in alveolar Type II epithelial cells and in subsets of
 CC nonciliated bronchiolar epithelial cells. TTF-1 was not detected in
 CC alveolar Type I cells or ciliated epithelial cells. In the adult lung,
 CC TTF-1 is detected in subsets of nonciliated bronchiolar epithelial cells
 CC and was most prominent in type II epithelial cells but was excluded from
 CC Type I cells.
 SQ Sequence 371 AA;
 R83015 Length: 371 February 11, 2000 15:48 Type: P Check: 5439 ..
 1 MSMSKHTTP FSVSDIISPL EESYKVGME GGGGLAPLAA YROGQAPPT
 51 AAMOOHAYGH HGAVTAAAYHM TAAGVPOLSH SAVGVCNGN LGMSLELPY
 101 ODTMRNSASG PGWYGANPDP REPATSRFMG PASGNNNSGM GGIISGIDVS
 151 KNAAPLEPAP RRRRYLFSQ AQVYELERRE KOQRYLSAPE REHLASMIHL
 201 TPPTQKIMFQ NHRRYKMKROA KDKAAQOOLQ QDSGGGGGGG GTGCGPQQQA
 251 QOOSPRRYAV PVLKDKGPC QAGAPAPGAA SLGCHAQQA CHQAQAQAA
 301 AAATVSGSG AGIGAPRHQ PGSAQSPDL AHHAAPAL QCVVSLSLH
 351 NSSGSDYGTM SCSTLLYGR T W
 !!AA_SEQUENCE 1.0
 ID W00725 standard; Protein: 207 AA.
 AC W00725;
 DT 30-NOV-1996 (first entry)
 DE Vascular endothelial growth factor-like protein SOM175.
 KM Vascular endothelial growth factor; VEGF; VEGF165; SOM175; neuron;
 OS astroglial proliferation.
 FH key Location/Qualifiers
 FT 1..21
 FT peptide /label= Sig-peptide
 FT W09627007-A1.
 PD 06-FEB-1996;
 PR 22-MAR-1995; AU-001457.
 PR 02-MAR-1995; AU-001457.
 PR 20-NOV-1995; AU-006647.
 PR 22-DEC-1995; AU-007274.
 PA (AMCA-) AMRAD OPERATIONS PTY LTD.
 PA Gilmond S. Hayward NK, Larsson C, Nordenskjold M;
 PI Webber G; 96-412774/41.
 DR WPI: 96-412774/41.
 DR N-PSDB: T33610.
 PT New growth factor related to vascular endothelial growth factor -
 PT useful for inducing astroglial proliferation and promoting neuronal

PT survival
 PS Claim 11; Page 41; 113pp; English.
 CC Human vascular endothelial growth factor (VEGF)-like polypeptide
 CC (W00725) is capable inducing the proliferation of vascular
 CC endothelial cells, of interacting with flt-1/FK1-1 receptors,
 CC and of inducing cell migration, cell survival and/or an increase
 CC in intracellular levels of alkaline phosphatase. It shows 33.3%
 CC identity with human VEGF (see also W00724). Splice variants
 CC (W00726-28) of SOM175 have also been identified. Recombinant SOM175
 CC can be produced in host cells transfected with vectors carrying
 CC SOM175 cDNA (see also T33610). It is useful for inducing astroglial
 CC proliferation and for promoting neural survival and/or proliferation.
 SQ Sequence 207 AA;
 W00725 Length: 207 February 11, 2000 15:48 Type: P Check: 1679 ..
 1 MSPLRLRL AALLQAPAQ AVSQPDAPG HQKRVSMID VITRATQPR
 51 EYVPELTEL NGYAKQLVP SCVTQRCGG CCPDGLCEV PICQHYRQ
 101 ILMIRPSSQ LGEMSLEHS OECERPCKD SAVPDAAT PHHROPKSV
 151 PWDAPAPGAP SPADINHTP ACPGSAHAP STTSALTPG AAAADAANS
 201 SVAKGGA
 !!AA_SEQUENCE 1.0
 ID R99253 standard; Protein: 374 AA.
 AC R99253;
 DT 05-DEC-1996 (first entry)
 DE Cytoplasmic antiprotease-2 protein.
 DE Cytoplasmic antiprotease-2 protein; CAP-2; serpin;
 KW serine protease inhibitor; antinflammatory; apoptosis.
 OS Homo sapiens.
 FH key Location/Qualifiers
 FT modified_site 8
 FT modified_site /label= N-glycosylation_site
 FT modified_site 78
 FT modified_site /label= N-glycosylation_site
 FT W09624650-A2.
 PD 15-AUG-1996;
 PD 02-FEB-1996; U01288.
 PR 08-FEB-1995; US-385500.
 PA (ZMO) ZYMOGENETICS INC.
 PI Sprecher CA;
 DR WPI: 96-393014/39.
 DR N-PSDB: T35220.
 PT Human cytoplasmic antiprotease-2 (CAP-2) and CAP-3 - serine
 PT protease inhibitors useful in the purification of proteins and in
 PT the treatment of inflammatory diseases and apoptosis
 PS Claim 9; Page 37-38; 50pp; English.
 CC Human cytoplasmic antiprotease-2 protein (CAP-2) (R99253) is an
 CC inhibitor of specific trypsin-like serine proteases. Its amino
 CC acid sequence was deduced from a cDNA clone (135220) isolated from
 CC a human placenta lambda gt11 cDNA library. Another serpin, CAP-3
 CC (R99254), has also been identified. The CAPs are useful for
 CC the purification of proteins and in the treatment of inflammatory
 CC diseases and apoptosis. They can be expressed in transformed or
 CC transfected host (pref. mammalian) cells.
 SQ Sequence 374 AA;
 R99253 Length: 374 February 11, 2000 15:48 Type: P Check: 4145 ..
 1 MDLCEANGT FALFEKILG EEDNSRNVEF SPMSSSALA WYMGAKGST
 51 AAMSOALCL YKGDHHRGF OSLSSEVNT GTQYLRTAN RLRGKTCDF
 101 LDFEYECOR FYQAELELS FAEDTECKR HINWVAEXT EGRISVDA
 151 GTVDPLRLV LVAIFYEKG WNEQFDKRYT RGLFKTNEE KTYQWMEKE
 201 AKFKGYADE VHTQVLELPY VEELSNTL LPDQNTDLAV VKALTYEKF

```

251 KANTSEKIT KSKYQVLPFR LKLESYDLE PFLRLGMD APDEKADFS
301 GMSTERNVPL SKVAHKCEVE VNEEGTEAAL ATAVRSRSC SRMBEPCAD
351 HPFLFIRRH KTNCLIFGCR FSSP

!!AA SEQUENCE 1.0
ID W01743 standard; Protein: 259 AA.
AC W01743;
DE 22-APR-1997 (first entry)
DE C. pneumoniae 53 kDa antigen.
DE antigen; antibody; detection; determination; epitope.
OS Chlamydia pneumoniae.
PN J08304404-A.
PD 22-NOV-1996.
PR 28-APR-1995; JP-106014.
PR (HTB ) HITACHI CHEM CO LTD.
PR WPI: 97-056178/06.
DR N-PSDB: T59311.
DE Detection and determination of anti-Chlamydia pneumoniae antibody -
PI using the polypeptide C as the antigen
PS Example 3; Page 17-18; 18pp; Japanese.
CC The protein is that of the Chlamydia pneumoniae 53 kDa antigen. A
CC method for the detection and determination of anti-C. pneumoniae
CC antibodies in a sample comprises using at least 5 consecutive amino
CC acids of the polypeptide C 73 kDa antigen (W01742).
SQ Sequence 259 AA.

W01743 Length: 259 February 11, 2000 15:48 Type: P Check: 9728

1 MSISSSGPD NQNMISOVL TSTPGVPOQ DKLSGNTRQ IQOTROGKNT
51 EMESDATNG ASGDKTSST TKTERAPOQ VAAKESSES QKAGADTVS
101 GAATATSNF ATKIMQTSI EEMSKSMST LESQUSAA QKKEEAVVV
151 AALSGKSGS AKLEPELPK PGVTRSEVI EIGALAKAI QRLGATKSA
201 LSNVASTQAO ADQNKGLGE KQAKIKDKER EYEGEMAAE QSKDLEGIM
251 DTVNTVMAA

!!AA SEQUENCE 1.0
ID W04831 standard; Protein: 207 AA.
AC W04831;
DE 28-APR-1997 (first entry)
DE Vascular endothelial growth factor-B186.
DE Vascular endothelial growth factor; VEGF;
DE VEGF; endothelium; mesodermal cell; cationic dimer; tissue regeneration;
DE VEGF; vascular permeability factor; cell mitogen; angiogenesis; cell growth;
DE embryonic development; wound healing; tissue reorganization; antibody;
DE cancer; metastatic risk; tumour cell; mouse.
OS Homo sapiens.
PN W09626736-A1.
PD 06-SEP-1996.
PR 01-MAR-1996; U02957.
PR 01-MAR-1995; US-397651.
PR 06-JUN-1995; US-469427.
PR 06-DEC-1995; US-569063.
PR (LUDM-) ICDWIG INST CANCER RES.
PR (UHE-) UNIT HELSINKI LICENSING LTD OY.
PA WPI: 96-412582/41.
PA Aaltalo K; Eriksson U, Olofsson B, Pajusola K;
PI WPI: 96-412582/41.
DR N-PSDB: T37935.
DE Vascular endothelial growth factor VEGF-B proteins - useful to
DE accelerate angiogenesis in wound healing, also related nucleic acid
DE and antibodies for cancer diagnosis
PS Claim 18; Page 62; 107pp; English.
PS W04824-W04831 represent the vascular endothelial growth factor (VEGF)
CC proteins of the invention which promote endothelial or mesodermal cell
CC proliferation. VEGF is also a glycosylated cationic dimer, and is
CC sometimes referred to as vascular permeability factor (VPF). VEGF has

```

```

CC diverse effects, depending on the specific biological context in which it
CC is found. VEGF is a potent endothelial cell mitogen, and directly
CC contributes to induction of angiogenesis in vivo by promoting endothelial
CC cell growth during normal embryonic development, wound healing, and
CC tissue regeneration/reorganization. The VEGF proteins of the invention
CC share the angiogenic and other properties of VEGF, but are distributed
CC and expressed in tissues differently to VEGF. The proteins can therefore
CC be used to accelerate angiogenesis in wound healing. Antibodies against
CC the proteins can be used for inhibiting angiogenesis. The antibodies can
CC also be used diagnostically to quantitatively detect VEGF-B. Primers
CC complementary to the coding sequences for the proteins of the invention
CC can also be used to detect VEGF-B coding sequences. Quantification of
CC VEGF-B in cancer biopsy specimens may be useful as an indicator of
CC metastatic risk. VEGF-B expression in a cell can be retarded using
CC antisense sequences directed against the VEGF coding sequences, this is
CC especially useful in retarding VEGF expression in tumour cells.
SQ Sequence 207 AA.

W04831 Length: 207 February 11, 2000 15:48 Type: P Check: 1679

1 MSPLRLRL ALIQLAPQ APVSQDPAG HQKRVSMID VYTRATCPFR
51 EYVYPLVLEI MGTVAKQLVP SCYTVQRCGG CCPDDGLECV PTGOHVRMQ
101 IIMIRYPSO IGEMSLSEHS QCECRPKKD SAVKPDRAAT PHHRPOPRSV
151 PGWDSAPGAP SPADITHTP APGPSAHAAP STSALITGP AAAADDAAS
201 SVAKGGA

!!AA SEQUENCE 1.0
ID W11481 standard; Protein: 205 AA.
AC W11481;
DE 22-APR-1997 (first entry)
DE D. immitis mature venom allergen antigen 5-1-like protein PDIVA205.
DE Venom allergen antigen 5-1-like gene; VAS; helminth; parasite;
DE PDIVA615; PDIVA205; vaccine; heartworm.
OS Dirofilaria immitis.
PN W096337218-A1.
PD 28-NOV-1996.
PR 23-MAY-1996; U07709.
PR 23-MAY-1995; US-450944.
PR (HESK-) HESKA CORP.
PA WPI: 97-020935/02.
PA TRIP CA Wisconsin N;
PI WPI: 97-020935/02.
DR N-PSDB: T51379.
DE Nucleic acid encoding helminth venom allergen antigen 5-1-like protein
DE - pref. from Dirofilaria immitis or Onchocerca volvulus, useful esp.
DE in vaccines to prevent helminth infection
PS Claim 19; Page 9192; 124pp; English.
PS Dirofilaria immitis mature venom allergen antigen 5-1-like protein
CC VAS (W11481), or PDIVA205, is capable of eliciting an immune
CC response (cellular and/or humoral) in animals. It is encoded by a
CC cDNA molecule ndIVA615 (T51379). VAS-1-like proteins (see also
CC W11480-86) can be used in vaccines to prevent helminth infection.
SQ Sequence 205 AA.

W11481 Length: 205 February 11, 2000 15:48 Type: P Check: 5195

1 YECEGGKILP TERKNIVTOI NKYSRLING KKKKKDYLK PGKNNLRLMR
51 WDCKLEKSAQ MNWANNCEVGH SPSSERRGIG ENYAYVSSG SVRDLKKTNG
101 TDAGRLWMSL LERKYSNDPS NNLTSVAME NILHFTQMAW GELYKGSQV
151 DHNIVAVART LVFICHYPPG GNNVKDLIYE LGNPKKHND CRTKRSASAS
201 GLCKK

!!AA SEQUENCE 1.0
ID W11479 standard; Protein: 221 AA.
AC W11479;
DE 22-APR-1997 (first entry)

```

DE D. limittis venom allergen antigen 5-like protein PDIVA221.
 KW Venom allergen antigen 5-like gene; VAS; helminth; parasite;
 OS PDIVA221; vaccine; heartworm.
 OS Dicrofilaria limittis.
 FT Key location/Qualifiers
 FT peptide 1..16
 FT 17..221 Sig-peptide
 FT protein 17..221
 FT /label=Sig-peptide
 FT /note= mature protein is represented by PDIVA205"
 PN W069327218-A1.
 PD 28-NOV-1996.
 PE 23-MAY-1996; U07709.
 PR 23-MAY-1995; US-450944.
 PA (HEK-) HESKA CORP.
 PI TILIP 2A, Wisniewski N;
 DR WPI 97-020935/02.
 DR N-PSDB; T51376.
 DR N-PSDB; T51376.
 PT Nucleic acid encoding helminth venom allergen antigen 5-like protein
 PT in vaccines to prevent helminth infection
 PS Clam 13; Page 89-90; 124pp; English.
 CC Dicrofilaria limittis venom allergen antigen 5-like protein VAS
 CC (W11479), or PDIVA221, is capable of eliciting an immune response
 CC (cellular and/or humoral) in animals. It is encoded by a cDNA
 CC clone (T51376) isolated from a D. limittis cDNA library by screening
 CC with immune dog serum. VAS-like proteins (see also W11480-86) can
 CC be used in vaccines to prevent helminth infection.
 SQ Sequence 221 AA;
 W11479 Length: 221 February 11, 2000 15:48 Type: P Check: 9389 ..
 1 MILVIFPAI IVVASYEC GSKLPTERK NIVQINKYR SLIRKIKLN
 51 KDGYLMPGK NMLRMMDCK LEKSNQNMN MCYFGHSPSS ERKIGENVY
 101 AYMSGSVDR LKKTACTDAG RLWSELEKLY YSNPNSNLT SEVAMENILH
 151 FTQWAGETV KLGSGVDHNI VMVARTLVTI CHYFGGNNV KDLYELGNP
 201 CRHKKDCKTR RCSAKSGCLK K
 11AA SEQUENCE 1.0
 ID W13272 standard; Protein: 244 AA.
 AC W13272;
 DT 02-JUN-1997 (first entry)
 DE Rhodococcus erythropolis SK92-B1 regulatory factor component.
 KW Regulatory factor; activation; nitrilase; gene promoter;
 KW recombinant plasmid; transformation; microorganism; catalyst;
 OS Rhodococcus erythropolis.
 PN J09028382A.
 PD 04-FEB-1997.
 PE 19-JUL-1995; 204061.
 PR 19-JUL-1995; JP-204061.
 DR (NITR) NITRO CHEM IND CO LTD.
 DR WPI: 97-159092/15.
 DR N-PSDB; T46931.
 PT Microorganism producing nitrilase in absence of inducer - useful as
 PT catalyst in large scale organic acid production
 PS Clam 7; Pages 7-8; 13pp; Japanese.
 CC The present sequence is one of the two components of the
 CC Rhodococcus erythropolis SK92-B1 regulatory factor (RF), which
 CC activates a nitrilase gene promoter. A recombinant plasmid
 CC containing the RF gene can be used to transform a microorganism,
 CC enabling it to produce nitrilase in high yield without adding any
 CC inducer to the medium. Hence, the preparation of the medium is easy
 CC for industrial large scale culture of the microorganism, and the
 CC disposal cost of the waste culture is reduced. Nitrilase is
 CC industrially useful as a catalyst for the production of organic
 CC acids, as it hydrolyses nitriles into their corresponding acids.
 SQ Sequence 244 AA;

W13272 Length: 244 February 11, 2000 15:48 Type: P Check: 2018 ..
 1 MAGADVHAOG GTNRARILV VDDEKVRMT VTWGLESENF DYAADADGA
 51 ALROYTESAP DLWVLDLSP GKGLEVLAT VRTRDALPIV VLTARDETE
 101 RIVALDLGAD DYVKRPSFR ELARIRAVL RRTAEPRHE AAVORFGLD
 151 IDTAREVRL HGIPLEFTK EFDLAYMA SPWQVSRRR LLEVRSSP
 201 DWQDATVTE HVHRIRKIE EDPFTILO TVAGAGYRD GERA
 11AA SEQUENCE 1.0
 ID W06943 standard; Protein: 140 AA.
 AC W06943;
 DT 30-JUN-1997 (first entry)
 DE CagI locus product 14.
 KW CagI; CagA; virulence factor; exporter molecule; homology; pti gene;
 KW Bordetella pertussis; VIR B4; Agrobacterium tumefaciens; invasion factor;
 KW Salmoneella; type I strain; virulence; diagnosis; H. pylori; infection;
 KW vaccine; treatment; duodenal; gastric ulcer; active gastritis;
 KW adenocarcinoma.
 OS Helicobacter pylori.
 PN W06933274-A1.
 PD 24-OCT-1996.
 PE 18-APR-1996; IB0343.
 PR 20-APR-1995; US-425184.
 PR 07-JUN-1995; US-477451.
 PA (BIOC-) BIOCINE SPA.
 PI Covacci A;
 DR WPI: 96-485780/48.
 DR N-PSDB; T46159.
 PT Helicobacter pylori CagI polynucleotide and related proteins - used
 PT in diagnosis and in vaccines for the treatment of H. pylori
 PT infection associated disease
 PS Disclosure; Fig 41; 303pp; English.
 CC The present sequence is encoded by a putative open reading frame (ORF) of
 CC the CagI locus (see T46159). The CagI region contains clusters of
 CC putative ORFs with different polarities. It is hypothesized that some of
 CC the ORFs may encode exporter molecules with homology to the pti genes of
 CC Bordetella pertussis and VIR B4 genes of Agrobacterium tumefaciens and
 CC for proteins with motifs shared by the purported invasion factors of
 CC Salmoneella genus. The absence of the CagA gene in the type I strains is
 CC associated with the absence of CagI sequences (which may encode virulence
 CC factors restricted to type I strains). The CagI nucleotide sequence, its
 CC fragments and encoded proteins are used in the diagnosis of H. pylori
 CC (esp. H. pylori type I strain) infection in an individual and in vaccines
 CC (claimed) for the treatment of H. pylori infection associated with e.g.
 CC duodenal and gastric ulcers, severe forms of active gastritis (esp. type
 CC Sequence 140 AA;
 W06943 Length: 140 February 11, 2000 15:48 Type: P Check: 9316 ..
 1 KRIFIKNTY LGLVSLACL QSVNADONT INDISPEDMA LNSGLVSRD
 51 QLKIEIPKER LEQKAILND YNDKNNVNF DDISGSFOP NDNGLINAMW
 101 GIGNLMSQW MDGYPPNPF MYGAYPTYS D SSFLPILIGY
 11AA SEQUENCE 1.0
 ID W06935 standard; Protein: 136 AA.
 AC W06935;
 DT 30-JUN-1997 (first entry)
 DE CagI locus product 6.
 KW CagI; CagA; virulence factor; exporter molecule; homology; pti gene;
 KW Bordetella pertussis; VIR B4; Agrobacterium tumefaciens; invasion factor;
 KW Salmoneella; type I strain; virulence; diagnosis; H. pylori; infection;
 KW vaccine; treatment; duodenal; gastric ulcer; active gastritis;
 KW adenocarcinoma.
 OS Helicobacter pylori.
 PN W06933274-A1.

PD 24-OCT-1996.
 PF 18-APR-1996; IB0343.
 PR 20-APR-1995; US-425194.
 PA (BIOC-) BIOCINE SPA.
 PI Covacci A.
 DR WPI: 96-485780/48.
 DR N-PSDB: T46159.
 PI Helicobacter pylori CagI polynucleotide and related proteins - used
 in diagnosis and in vaccines for the treatment of H. pylori
 infection associated disease
 PI Disclosure: Fig 4D: 303pp; English.
 PS The present sequence is encoded by a putative open reading frame (ORF) of
 the CagI locus (see T46159). The CagI region contains clusters of
 putative ORFs with different polarities. It is hypothesized that some of
 the ORFs may encode exporter molecules with homology to the Pil genes of
 Bordetella pertussis and VIR B4 genes of Agrobacterium tumefaciens and
 for proteins with motifs shared by the putative invasion factors of
 Salmonella genus. The absence of CagI sequences (which may encode virulence
 factors) restricted to type I strains). The CagI nucleotide sequence, its
 fragments and encoded proteins are used in the diagnosis of H. pylori
 (esp. H. pylori type I strain) infection in an individual and in vaccines
 (claimed) for the treatment of H. pylori infection associated with e.g.
 duodenal and gastric ulcers, severe forms of active gastritis (esp. type
 gastritis) and gastric adenocarcinoma.
 CC Sequence 136 AA;
 SQ
 W06935 Length: 136 February 11, 2000 15:48 Type: P Check: 8017 ..
 1 NIYFMALYK FTMALNFFKN QNGTNOISKL KQNFLOFKYS FNKHLDTYSL
 51 YRLEFNISST VIGFLIGFS YGAGVILVYP ILEFLALIK PSFYYTTYL
 101 ILLVSLSTIS KYLLSHAKF TKMLIMTQ MQWFL
 !!AA_SEQUENCE 1.0
 ID W20571 standard; Protein: 114 AA.
 AC W20571;
 DT 17-JUL-1997 (first entry)
 DE H. pylori secreted or periplasmic protein 80257.aa.
 KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
 KM binding compound; bacterium; life cycle; activator; bacterium; inhibitor;
 KM duodenal ulcer disease; chronic gastritis; diagnosis; envelope.
 OS Helicobacter pylori.
 PN W09640893-A1.
 PD 19-DEC-1996.
 PF 06-JUN-1996; U09122.
 PR 07-JUN-1995; US-487032.
 PA (ASTR-) ASTRA AB.
 PI Berglindh OT, Smith D, Mellgaard BL;
 DR WPI: 97-052306/05.
 DR N-PSDB: T67714.
 PI Helicobacter pylori nucleic acid sequences and related
 polypeptide(s) - useful for vaccines to treat or prevent H. pylori
 infection.
 PS Claim 72: page 725-726; 1481pp; English.
 CC This sequence is a H. pylori secreted or periplasmic protein.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 infection or to identify H. pylori polypeptide binding compounds.
 CC Useful as potential H. pylori life cycle activators or inhibitors.
 CC The genomic contigs generated by mechanically shearing the bacterial
 DNA. The sequences were analysed for ORF of at least 180 nucleotides,
 and the predicted coding regions defined by computer evaluation. To
 identify likely H. pylori antigens for vaccine development, the amino
 acid sequences predicted from various ORF were analysed for significant
 homology to other known or exported membrane proteins. Having identified
 CC and determined the sequences of interest, particular regions can be
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide
 production, e.g. in E. coli hosts.
 CC Sequence 114 AA;
 SQ

W20571 Length: 114 February 11, 2000 15:48 Type: P Check: 6963 ..
 1 MKRPISKLMQ NLFQKHSFN KALDRYSLEY RLNFSSIVY GFLALSYSG
 51 AGVILVYPL FLFLALIKPS FYYTTYLL LVSLISISK YLLSHANTM
 101 KLIMTQMQ NMFL
 !!AA_SEQUENCE 1.0
 ID W20886 standard; Protein: 154 AA.
 AC W20886;
 DT 18-JUL-1997 (first entry)
 DE H. pylori secreted or periplasmic protein, 13ee12016orf8.
 KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
 KM identification; binding compound; bacterium; life cycle; activator;
 KM bacteria; inhibitor; duodenal ulcer disease; chronic gastritis;
 diagnosis.
 OS Helicobacter pylori.
 PN W09640893-A1.
 PD 19-DEC-1996.
 PF 06-JUN-1996; U09122.
 PR 07-JUN-1995; US-487032.
 PA (ASTR-) ASTRA AB.
 PI Berglindh OT, Smith D, Mellgaard BL;
 DR WPI: 97-052306/05.
 DR N-PSDB: T68139.
 PI Helicobacter pylori nucleic acid sequences and related
 polypeptide(s) - useful for vaccines to treat or prevent H. pylori
 infection, and to detect Helicobacter
 PS Claim 72: page 1289; 1481pp; English.
 CC This sequence represents a H. pylori secreted or periplasmic protein.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 infection or to identify H. pylori polypeptide binding compounds.
 CC Useful as potential H. pylori life cycle activators or inhibitors.
 CC The genomic sequence of H. pylori (ATCC 35619) was determined from
 overlapping contigs generated by mechanically shearing the bacterial
 DNA. The sequences were analysed for ORF of at least 180 nucleotides,
 and the predicted coding regions defined by computer evaluation. To
 identify likely H. pylori antigens for vaccine development, the amino
 acid sequences predicted from various ORF were analysed for significant
 homology to other known or exported membrane proteins. Having identified
 CC and determined the sequences of interest, particular regions can be
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide
 production, e.g. in E. coli hosts.
 CC Sequence 154 AA;
 SQ
 W20886 Length: 154 February 11, 2000 15:48 Type: P Check: 2872 ..
 1 NNAKKNKGV KSKMTFFYKI KLFWACLLI GMFNAPLNAV QNTDIDISP
 51 EDMAINSYGL VSRDQAKIEI PKETLEQKVA IINDYNDKNV NIKEDDISLG
 101 SFQPDNLTGI NAMWGIONLL MSQMSNYPG NNSFMVGYAP TYSDSFLPP
 151 IIGY
 !!AA_SEQUENCE 1.0
 ID W20776 standard; Protein: 276 AA.
 AC W20776;
 DT 15-JUL-1997 (first entry)
 DE H. pylori flagella-associated protein, 07ge20415orf34.
 KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
 KM identification; binding compound; bacterium; life cycle; activator;
 KM bacteria; inhibitor; duodenal ulcer disease; chronic gastritis;
 diagnosis.
 OS Helicobacter pylori.
 PN W09640893-A1.
 PD 19-DEC-1996.
 PF 06-JUN-1996; U09122.
 PR 07-JUN-1995; US-487032.
 PA (ASTR-) ASTRA AB.
 PI Berglindh OT, Smith D, Mellgaard BL;
 DR WPI: 97-052306/05.
 DR N-PSDB: T68139.
 PI Helicobacter pylori nucleic acid sequences and related
 polypeptide(s) - useful for vaccines to treat or prevent H. pylori
 infection, and to detect Helicobacter
 PS Claim 72: page 1289; 1481pp; English.
 CC This sequence represents a H. pylori secreted or periplasmic protein.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 infection or to identify H. pylori polypeptide binding compounds.
 CC Useful as potential H. pylori life cycle activators or inhibitors.
 CC The genomic sequence of H. pylori (ATCC 35619) was determined from
 overlapping contigs generated by mechanically shearing the bacterial
 DNA. The sequences were analysed for ORF of at least 180 nucleotides,
 and the predicted coding regions defined by computer evaluation. To
 identify likely H. pylori antigens for vaccine development, the amino
 acid sequences predicted from various ORF were analysed for significant
 homology to other known or exported membrane proteins. Having identified
 CC and determined the sequences of interest, particular regions can be
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide
 production, e.g. in E. coli hosts.
 CC Sequence 276 AA;
 SQ

PA (ASTR) ASTRA AB.
 PI Berglundh OT, Smith D, Mellgaard BL;
 DR WPI: 97-052306/05.
 DR N-PSDB: T674029.
 PT Helicobacter pylori nucleic acid sequences and related
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
 PS infection, and to detect Helicobacter
 PS Claim 56; Page 1187; 1481pp; English.
 CC This sequence represents a H. pylori flagella-associated protein.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori 11fe cycle activators or inhibitors.
 CC The genomic sequence of H. pylori 11fe cycle activators or inhibitors
 CC overlapping contigs generated by mechanically shearing the bacterial
 CC DNA. The sequences were analysed by computer evaluation. To
 CC identify likely H. pylori antigens for vaccine development, the amino
 CC acid sequences predicted from various ORF were analysed for significant
 CC homology to other known or exported membrane proteins. Having identified
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide
 CC production, e.g. in E. coli hosts.
 SQ Sequence 276 AA;
 W20776 Length: 276 February 11, 2000 15:48 Type: P Check: 6185 ..
 1 RUKMAYKLP KORAQDELIS MSEKAILILI QVEDTITGEI LRHDDISIT
 51 EISKOIVOLN GTDKQIGAAV LEEFALFOS NOYINNGGE YARELLITTL
 101 GSEARKVMD KLRKSLQTKR NFAVLGRKP QQLADITVE HPQTALILA
 151 HMEAPNAET LSYFPEMKA EISIRMANLG EISPOVYKRV STLENKLES
 201 LTSYKIEVGG LRAVAEIFNR LGQSAKTTL ARLESVDNKL AGAIKEMFT
 251 FEDIAKLDNF AIMRDFKSGG LKRTGL
 11AA-SEQUENCE 1.0
 ID W20259 standard; Protein: 108 AA.
 AC W20259;
 DT 09-JUL-1997 (first entry)
 DE H. pylori secreted or periplasmic protein 23564012.aa.
 KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;
 KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
 KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.
 OS Helicobacter pylori.
 PN W09640893-A1.
 PD 19-DEC-1996;
 PE 06-JUN-1996; U09122.
 PR 07-JUN-1995; US-487032.
 PR 01-APR-1996; US-630405.
 PA (ASTR) ASTRA AB.
 PI Berglundh OT, Smith D, Mellgaard BL;
 DR WPI: 97-052306/05.
 DR N-PSDB: T67473.
 PT Helicobacter pylori nucleic acid sequences and related
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
 PS infection, and to detect Helicobacter
 PS Claim 72; Page 463; 1481pp; English.
 CC This sequence is a H. pylori secreted or periplasmic protein.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds.
 CC The genomic sequence of H. pylori 11fe cycle activators or inhibitors.
 CC overlapping contigs generated by mechanically shearing the bacterial
 CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,
 CC and the predicted coding regions defined by computer evaluation. To
 CC identify likely H. pylori antigens for vaccine development, the amino
 CC acid sequences predicted from various ORF were analysed for significant
 CC homology to other known or exported membrane proteins. Having identified
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide

CC production, e.g. in E. coli hosts.
 SQ Sequence 108 AA;
 W20259 Length: 108 February 11, 2000 15:48 Type: P Check: 5141 ..
 1 MKTNYKRL LFAMCLISM FNAFLMDQN TDIKDISPD MALNSGLVS
 51 RDOKEIFPK ETEGEVTL NDYDKNVTI KDDDISLGSF QPNDLGINA
 101 MWGIONTL
 11AA-SEQUENCE 1.0
 ID W20324 standard; Protein: 102 AA.
 AC W20324;
 DT 09-JUL-1997 (first entry)
 DE H. pylori cell envelope transporter protein, 2461062.aa.
 KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;
 KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
 KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.
 OS Helicobacter pylori.
 PN W09640893-A1.
 PD 19-DEC-1996;
 PE 06-JUN-1996; U09122.
 PR 07-JUN-1995; US-487032.
 PR 01-APR-1996; US-630405.
 PA (ASTR) ASTRA AB.
 PI Berglundh OT, Smith D, Mellgaard BL;
 DR WPI: 97-052306/05.
 DR N-PSDB: T67528.
 PT Helicobacter pylori nucleic acid sequences and related
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
 PS infection, and to detect Helicobacter
 PS Claim 56; Pages 516; 1481pp; English.
 CC The present sequence is a Helicobacter pylori cell envelope protein.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds, useful
 CC as potential H. pylori 11fe cycle activators or inhibitors. The genomic
 CC sequence of H. pylori 11fe cycle activators or inhibitors. The genomic
 CC contigs generated by mechanically shearing the bacterial DNA. The
 CC sequences were analysed for ORF of at least 180 nucleotides, and the
 CC predicted coding regions defined by computer evaluation. To identify
 CC likely H. pylori antigens for vaccine development, the amino acid
 CC sequences predicted from various ORF were analysed for significant
 CC homology to other known or exported membrane proteins. Having identified
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide
 CC production, e.g. in E. coli hosts.
 SQ Sequence 102 AA;
 W20324 Length: 102 February 11, 2000 15:48 Type: P Check: 6372 ..
 1 MNAHSILVS KTSLSNLIIF VVQPOGKUSM TDAIDENMT NSGLRWYRVN
 51 EIAEKFKIKR DKALVYINK GCKNPLTKN YNIRNGEEL RVIKRPLPLVR
 101 DK
 11AA-SEQUENCE 1.0
 ID W20696 standard; Protein: 121 AA.
 AC W20696;
 DT 15-JUL-1997 (first entry)
 DE H. pylori secreted or periplasmic protein 05ae20220orf50.
 KW Secretion; vaccine; prevention; treatment; infection; identification;
 KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
 KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.
 OS Helicobacter pylori.
 PN W09640893-A1.
 PD 19-DEC-1996;
 PE 06-JUN-1996; U09122.
 PR 07-JUN-1995; US-487032.
 PR 01-APR-1996; US-630405.
 PA (ASTR) ASTRA AB.

Mon Feb 14 08:07:18 2000

ags.cat

Page 53

PI Berglindh OT, Smith D, Mellgaard BL;
DR WPI: 97-052306/05.
DR N-PSDB: 167949.
PT Helicobacter pylori nucleic acid sequences and related
PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
PS infection, and to detect Helicobacter
PS Claim 72: Page 1115; 1461pp; English.
CC The present sequence represents a Helicobacter pylori protein which
CC is likely to be secreted or passively absorbed.
CC The protein or to identify H. pylori polypeptide binding compounds,
CC infection or to identify H. pylori life cycle activators or inhibitors.
CC useful as potential H. pylori life cycle activators or inhibitors.
CC The genomic sequence of H. pylori (ATCC 55679) was determined from
CC overlapping contigs generated by mechanical shearing the bacterial
CC DNA. The sequences were analysed for ORF of at least 160 nucleotides,
CC and the predicted coding regions defined by computer evaluation. To
CC identify likely H. pylori antigens for vaccine development, the amino
CC acid sequences predicted from various ORF were analysed for significant
CC homology to other known or exported membrane proteins. Having identified
CC isolated from H. pylori by PCR amplification for recombinant polypeptide
CC production, e.g. in E. coli hosts.
SQ Sequence 121 AA;

W20696 Length: 121 February 11, 2000 15:48 Type: P Check: 7893 ..

1 TSSFTKMKR PISKMONL QPKSHFNKL DKSLYRLF NISSIVGFL
51 IALFSGAGV IIVYPIFLF ALIKPSFY YTYLLILVS LSIISKYLL
101 SHANTMKLI MMTQMNMF L

!!AA-SEQUENCE 1.0
ID W20139 standard; protein: 53 AA.
AC W20139;
DE 22-JUL-1997 (first entry)
DE H. pylori cytoplasmic protein, 14257751.ae.
KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;
KW binding compound; bacterium; life cycle; activator; bacteri; inhibitor;
KW duodenal; disease; chronic gastritis; diagnosis; envelope.
OS Helicobacter pylori.
PN W09640893.1.
PD 12-DEC-1996: U09122.
PF 06-JUN-1995: US-487032.
PR 01-JUN-1995: US-630405.
PA (ASTR) ASSTRA AB
PI Berglindh OT, Smith D, Mellgaard BL;
DR WPI: 97-052306/05.
DR N-PSDB: 167382.
PT Helicobacter pylori nucleic acid sequences and related
PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
PS infection, and to detect Helicobacter
PS Claim 71: Page 363; 1481pp; English.
CC The present sequence is a H. pylori cytoplasmic protein involved in
CC carbohydrate metabolism.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds,
CC infection or to identify H. pylori life cycle activators or inhibitors.
CC The genomic sequence of H. pylori (ATCC 55679) was determined from
CC overlapping contigs generated by mechanical shearing the bacterial
CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,
CC and the predicted coding regions defined by computer evaluation. To
CC identify likely H. pylori antigens for vaccine development, the amino
CC acid sequences predicted from various ORF were analysed for significant
CC homology to other known or exported membrane proteins. Having identified
CC isolated from H. pylori by PCR amplification for recombinant polypeptide
CC production, e.g. in E. coli hosts.
SQ Sequence 53 AA;

W20139 Length: 53 February 11, 2000 15:48 Type: P Check: 1367 ..

1 MIDPVLVIFG ATODLAMKRL FVSLYEIYIS FVWLKTIIGL SHRGVNSYPM
51 KSF
!!AA-SEQUENCE 1.0
ID W09433 standard; protein: 328 AA.
AC W09433;
DE 01-SEP-1997 (first entry)
DE Human placenta putative P-2u receptor, PNR.
DE PNR: putative P-2u receptor; placental; inflammation; diagnosis;
KW Carcinoma; neoplasia; cancer; cystic fibrosis; hypertension;
KW high blood pressure; infection.
OS Homo sapiens.
PD 05-DEC-1996.
PF 03-JUN-1995: U08481.
PR 02-JUN-1995: US-459046.
PA (INCR) INCYTE PHARM INC.
PI WPI: 97-042714/04.
DR N-PSDB: 147730.
PT Isolated human placenta putative P-2u receptor (PNR) gene
PT used to develop prods. for the diagnosis and treatment of conditions
PS associated with altered expression of the receptor e.g. inflammation
PS Claim 1: Page 27-29; 36pp; English.
CC W09433 shows a human placenta-derived putative P-2u receptor
CC designated PNR. P-2u is specifically expressed in cells active in
CC immunity. An assay for up-regulated expression of PNR can accelerate
CC diagnosis and proper treatment of conditions caused by abnormal signal
CC transduction due to systemic and local infections, traumatic and
CC other tissue damage, hereditary or environmental diseases associated
CC with hypertension, carcinomas, cystic fibrosis and other physiological
CC or pathological problems.
SQ Sequence 328 AA;

W09433 Length: 328 February 11, 2000 15:48 Type: P Check: 8653 ..

1 MEMDGTDOA LGIPPTCY RENFKDILP PYSAVLAPA LPLNICYITO
51 ICTSRALTR TAYTUNLAL PDLIACSLP LLIYNAOGD HMPGDFACR
101 LVRFLEFANL HGRILFLTCI SFORYLGICH PLAPMKRGG RRAAMLVCA
151 VWLAVTQCL PTAIFATCI QNRNTVCYL SPALATHYM PYGMALTVIG
201 FLPLFAALLA CYCLACRLC RODGPABVA QERRGRANM AVVVAANFGI
251 SFLPHITKT AYLAIVSTPG VPCVLEAFA AAYKGTTPFA SANSYLDPII
301 FYFOKKFRP RPHELOKIL DKMDROGR
!!AA-SEQUENCE 1.0
ID W21704 standard; protein: 248 AA.
AC W21704;
DE 26-SEP-1997 (first entry)
DE Luffin-A.
KW pro-Ribosome inactivating protein; proRIP; peptide linker; cancer;
KW inactivation; eukaryotic ribosome; alpha fragment; beta fragment;
KW inhibitor; protein synthesis; N-glycosidase; glycosidase; liver;
KW rct; ribosomal 28S RNA; cellular proliferation; HIV-infected T cell.
OS Luffa cylindrica.
FT Key
FT Location/Qualifiers
FT 138..148
FT /note="Position of possible insertion of internal
FT peptide linker sequence"
PD US563384-A.
PD 03-JUN-1997.
PF 11-JUN-1990: US-535636.
PR 11-JUN-1990: US-535636.
PR 09-DEC-1992: US-987927.
PR 26-JAN-1995: US-378761.
PA (DOMC) DOMELANCO.
PI Hey TD, Morgan AER, Walsh TA;

DR WPI: 97-309831/28.
 PT Inactive precursor of maize ribosome-inactivating protein - also
 PT chimeric ribosome-inactivating protein precursors containing
 PT Internal linker sequences
 PS Claim 2: Column 117-118; 121pp: English.
 CC The sequences given in W21698-710 represent Ribosome Inactivating
 CC Proteins (RIP's), which may be used in the construction of the
 CC Internal peptide linker. The precursor sequence is incapable of
 CC inactivating eukaryotic ribosomes, but can be converted by removal
 CC of the linker into a protein having alpha and beta fragments and being
 CC capable of inactivating eukaryotic ribosomes. RIPs are potent
 CC inhibitors of eukaryotic protein synthesis. They possess a highly
 CC specific N-glycosidase activity which cleaves the glycosidic bond of
 CC adenine 4324 of rat liver ribosomal 28S RNA. RIP's selectively inhibit
 CC cellular proliferation of cells, e.g. cancer cells and HIV-infected T
 CC cells. The inactive proRIP proteins make it possible to provide protein
 CC synthesis inhibitors with uses in practical and improved ways not before
 CC possible. The RIP can be used to make cytotoxic conjugates.
 SQ Sequence 248 AA;

W21704 Length: 248 February 11, 2000 15:48 Type: P Check: 7673 ..
 1 DVPFSLSGSS STYSKRTIGD LKRALPSNGT VYNLTLLSS ASGASRTTLM
 51 TLSNTDGKRI TVAVDSQLY IMGVLVNSTS YFNESDAKL ASQYFKGST
 101 IYTLPSYSGV EKLDTAAGKI REKIPLGKPA LDSALTTTFH YDSTAAAAF
 151 LVILDTAFA SRKTYIGQI IERISKNOVP SLATISLENS LMSALSKQIQ
 201 LAQTNNGAFK TPVYITDDKG QRVETITNVS KVTAKNIQL LNKONVA

!!AA_SEQUENCE 1.0
 ID W21709 standard; Protein: 250 AA.
 AC W21709 standard; Protein: 250 AA.
 DT 26-SEP-1997 (first entry)
 DE Luffin-B.
 KW pro-Ribosome Inactivating Protein; proRIP; peptide linker; cancer;
 KW inhibitor; protein synthesis; N-glycosidase; glycosidic bond; liver;
 KW Luffa cylindrica
 OS Luffa cylindrica
 FH Key location/Qualifiers
 FT 139-149
 FT /note="Position of possible insertion of Internal
 PN peptide linker sequence"

US5635384-A.
 PD 03-JUN-1997: 535636.
 PF 11-JUN-1990: US-535636.
 PR 09-DEC-1992: US-987927.
 PR 26-JAN-1995: US-378761.
 PA (DOMC) DOMELANCO.
 PI Hey TD, Morgan AER, Walsh TA;
 DR WPI: 97-309831/28.
 PT Inactive precursor of maize ribosome-inactivating protein - also
 PT chimeric ribosome-inactivating protein precursors containing
 PT Internal linker sequences
 PS Claim 2: Column 125-128; 121pp: English.
 CC The sequences given in W21698-710 represent Ribosome Inactivating
 CC Proteins (RIP's), which may be used in the construction of the
 CC Internal peptide linker. The precursor sequence is incapable of
 CC inactivating eukaryotic ribosomes, but can be converted by removal
 CC of the linker into a protein having alpha and beta fragments and being
 CC capable of inactivating eukaryotic ribosomes. RIPs are potent
 CC inhibitors of eukaryotic protein synthesis. They possess a highly
 CC specific N-glycosidase activity which cleaves the glycosidic bond of
 CC adenine 4324 of rat liver ribosomal 28S RNA. RIP's selectively inhibit
 CC cellular proliferation of cells e.g. cancer cells and HIV-infected T
 CC cells. The inactive proRIP proteins make it possible to provide protein
 CC synthesis inhibitors with uses in practical and improved ways not before

CC possible. The RIP can be used to make cytotoxic conjugates.
 SQ Sequence 250 AA;

W21709 Length: 250 February 11, 2000 15:48 Type: P Check: 7681 ..
 1 ANVSFSLSCA DSKSYKFT ALKRALPSKE KVSNIPLLP SASGARITL
 51 MOLSNDARA ITMAIDVTNV YIMGVLVNST SYFNESDAKL ASQYFKGS
 101 TVTLPSYSGV YERLQNAAGK IREKIPLGPR ALDSALTSIF HDSTAAAF
 151 FVILDTAFA ASREKTYEGQ IERIPKNEV PSPAALSLEN EAMSLSKQIQ
 201 QLAQTNNGAF RTPVYITDNK GQVETITNLA SKVQIKDVNS KLLNKONIA

!!AA_SEQUENCE 1.0
 ID W10031 standard; Protein: 74 AA.
 AC W10031 standard; Protein: 74 AA.
 DT 24-SEP-1997 (first entry)
 DE Protein encoded by clone G-4-5 from activated maize embryo.
 KW Primer: polymerase chain reaction; cDNA library; anchored end; PCSUB;
 KW lock-docking; screening; PCR-based cDNA subtractive cloning; zea mays;
 KW maize; fusarium moniliforme.
 OS W0960038-A1.
 PD 19-DEC-1996.
 PF 05-JUN-1996: U08582.
 PR 07-JUN-1995: US-481687.
 PA (PRON-) BIONEER HARBED INT INC.
 PI Bt1995 SB DUNYACK JP, Wang X;
 DR WPI:97-067067/08.
 DR N-BSDP: T70121.
 PT Method for Proch. of cDNA libraries with anchored ends - useful for
 PT subtractive cloning of sequences of interest
 PS This protein is the longest putative.
 CC clone G-4-5 which was activated upon Fusarium moniliforme infection of
 CC germinating library with anchored ends using a method of isolation
 CC designated ICEPAL (isolation of cDNA ends from an anchored library). The
 CC method of generating a cDNA library with anchored ends produces discrete
 CC sized PCR products which do not necessarily require further subcloning/
 CC screening. The method also produces full-length cDNA's obtainable from
 CC the libraries as opposed to specific cDNA clones, as produced by
 CC previously known methods. Other methods such as PCR and RACE require a
 CC knowledge of the target sequence to be amplified, by using the PCSUB
 SQ method no previous knowledge is necessary.
 Sequence 74 AA;

W10031 Length: 74 February 11, 2000 15:48 Type: P Check: 8373 ..

1 WILRLRPHVH PPGASKEEP GRPGGRRRR RQTVLLASIV LASLPIETIV
 51 CMCGYCMHLP TLTVVHXYAA SCRP

!!AA_SEQUENCE 1.0
 ID W13647 standard; peptide: 30 AA.
 AC W13647.
 DT 25-NOV-1997 (first entry)
 DE Invariant chain region which contains the CLIP sequence.
 KW Class II associated invariant chain peptide; CLIP region; MHC;
 KW major histocompatibility complex; T cell epitope; antigen processing;
 KW p35 isoform; vaccine; autoimmune disorder; rheumatoid arthritis;
 KW multiple sclerosis; insulin dependent diabetes mellitus;
 KW myasthenia gravis; allergy; tumor; virus; parasite; mycobacterium;
 KW infection; transplant rejection; immunostimulant; diagnosis.
 OS Homo sapiens.
 PD W09710335-A1.
 PF 16-SEP-1996: D01763.
 PR 14-MAY-1996: US-017663.
 PR 14-SEP-1995: DE-034170.
 PA (MELM/) MELMS A.

PA (UYTU-) UNIV TUEBINGEN.
 PT Matcherek G, Meilms A.
 DR WPI: 97-202231/18.
 PT Recombinant polypeptide useful as immunostimulant and diagnostic
 tool - based on primary sequence of invariant chain contg. primary
 sequence of specific T-cell epitope
 PS Disclosure: Page 8; 29pp; German.
 CC This is the sequence of amino acids 97-126 from the p35 isoform of
 the invariant chain (i.e. of a polypeptide which associates with
 CC major histocompatibility complex class II molecules in the
 CC endoplasmic reticulum to form a class II loading compartment).
 CC Claimed polypeptides are derived from the invariant chain by
 CC replacing at least part of the CLIP (Class II associated invariant
 CC chain peptide) sequence with a T cell epitope. The recombinant
 CC polypeptide is processed in antigen presenting cells to produce the
 CC specific T cell epitope. Such polypeptides are useful in vaccines
 CC for the treatment of autoimmune disorders, e.g. rheumatoid arthritis,
 CC multiple sclerosis, insulin-dependent diabetes mellitus and myasthenia
 CC gravis, allergies and tumours, for the control of viral, parasitic and
 CC mycobacterial infections, and for use against transplant rejections.
 CC They can also be used in the diagnosis of autoimmune disorders.
 CC Vaccines produced using the recombinant polypeptides have a higher
 CC immunostimulant specificity and selectivity over prior art vaccines.
 SQ Sequence 30 AA;
 W13647 Length: 30 February 11, 2000 15:48 Type: P Check: 5362 ..
 1 LPRPKPVSK MRMATPLMQ ALPMGALPOG
 11AA_SEQUENCE 1.0
 ID W25141 standard; Protein: 248 AA.
 AC W25141;
 DT 02-DEC-1997 (first entry)
 DE Luffin-A (a ribosome inhibitory protein) inactive precursor.
 KW Maize; PRO-RIP; ribosome inactivating protein; alpha; beta subunit;
 KW Internal linker; Barley translation inhibitor; Trichosanthin;
 KW Ricin A-chain; Abrin-A-chain; Sapotin; SLR-1; Luffin A; MAP;
 KW Ricinus communis agglutinin; Momordin; PAP-S; Luffin-B; Dianthin 30;
 KW therapeutic toxin; tumour cell targeted; protein synthesis inhibitor;
 KW post-translational modification; cancer; neoplasia; HIV; AIDS;
 KW human immunodeficiency virus; acquired immune deficiency syndrome.
 OS Synthetic.
 PN US5646026-A.
 PD 08-JUL-1997.
 PF 11-JUN-1990; 535636.
 PR 09-DEC-1992; US-987927.
 PR 11-JUN-1990; US-535636.
 PR 26-JAN-1995; US-378761.
 PR 07-JUN-1995; US-485286.
 PA (DOMC) DOMELANCO.
 PI Hey TD, Morgan AER, Walsh TA;
 DR WPI: 97-362934/33.
 DT DNA encoding pro-ribosome inactivating proteins - inactive
 PT precursors of ribosome inactivating proteins; can be expressed in
 PT eukaryotic cells without causing cell death
 PS Claim 4; Column 11-120; 186pp; English.
 CC W25141 shows a Luffin-A (ribosome inhibitory protein, RIP) protein
 CC which was engineered to contain a selectively removable internal peptide
 CC linker sequence separating the alpha and beta units of the RIP. When
 CC separated the two units regain activity and are capable of inactivating
 CC eukaryotic ribosomes and hence preventing protein production. Many
 CC different RIPs may be produced with an internal linker including
 CC maize RIP, Trichosanthin, Ricin A-chain, Abrin-A-chain and
 CC Sapotin. The RIPs can be used in the construction of therapeutic
 CC toxins targeted to specific cells such as tumour cells via the
 CC attachment of a targeting polypeptide, e.g. a monoclonal antibody.
 CC A further use is in HIV therapy (see US469903). There is interest
 CC in expressing RIP recombinantly in host eukaryotic cells, because of
 CC the capacity to provide correct post-translational processing. However,
 CC RIPs effectively inhibit protein synthesis in eukaryotic cells resulting
 CC in cell death. Since the inactive RIP proteins are not cytotoxic to
 CC eukaryotic cells, they can be recombinantly expressed in such cells and
 CC then converted to active RIP proteins.

SQ Sequence 248 AA;
 W25141 Length: 248 February 11, 2000 15:48 Type: P Check: 7673 ..
 1 DVRESLSSS STSYSFID LKRALPSNGT VYMLTILIS ASGASRYTLM
 51 TISNYDGAIT TVADVSOQL IMGYLVNST YFNESDAKI ASQYFKGST
 101 IYTLPSGNY EKLQTAGKI REKIPLGPA LDSALTTIIF YDSTAAAF
 151 IYLIQTFAE SRREYEGOI IERISKNOV SLATILENS LWSLSKOIO
 201 IAOINNGTFR TPVYITDCK QVEITVNTS KVTYKNIQL LNYKONVA
 11AA_SEQUENCE 1.0
 ID W25146 standard; Protein: 250 AA.
 AC W25146;
 DT 02-DEC-1997 (first entry)
 DE Luffin-B (a ribosome inhibitory protein) inactive precursor.
 KW Maize; PRO-RIP; ribosome inactivating protein; alpha; beta subunit;
 KW Internal linker; Barley translation inhibitor; Trichosanthin;
 KW Ricin A-chain; Abrin-A-chain; Sapotin; SLR-1; Luffin A; MAP;
 KW Ricinus communis agglutinin; Momordin; PAP-S; Luffin-B; Dianthin 30;
 KW therapeutic toxin; tumour cell targeted; protein synthesis inhibitor;
 KW post-translational modification; cancer; neoplasia; HIV; AIDS;
 KW human immunodeficiency virus; acquired immune deficiency syndrome.
 OS Synthetic.
 PN US5646026-A.
 PD 08-JUL-1997.
 PF 11-JUN-1990; 535636.
 PR 09-DEC-1992; US-987927.
 PR 11-JUN-1990; US-535636.
 PR 26-JAN-1995; US-378761.
 PR 07-JUN-1995; US-485286.
 PA (DOMC) DOMELANCO.
 PI Hey TD, Morgan AER, Walsh TA;
 DR WPI: 97-362934/33.
 DT DNA encoding pro-ribosome inactivating proteins - inactive
 PT precursors of ribosome inactivating proteins; can be expressed in
 PT eukaryotic cells without causing cell death
 PS Claim 4; Column 127-128; 186pp; English.
 CC W25146 shows a Luffin-B (ribosome inhibitory protein, RIP) protein
 CC which was engineered to contain a selectively removable internal peptide
 CC linker sequence separating the alpha and beta units of the RIP. When
 CC separated the two units regain activity and are capable of inactivating
 CC eukaryotic ribosomes and hence preventing protein production. Many
 CC different RIPs may be produced with an internal linker including
 CC maize RIP, Trichosanthin, Ricin A-chain, Abrin-A-chain and
 CC Sapotin. The RIPs can be used in the construction of therapeutic
 CC toxins targeted to specific cells such as tumour cells via the
 CC attachment of a targeting polypeptide, e.g. a monoclonal antibody.
 CC A further use is in HIV therapy (see US469903). There is interest
 CC in expressing RIP recombinantly in host eukaryotic cells, because of
 CC the capacity to provide correct post-translational processing. However,
 CC RIPs effectively inhibit protein synthesis in eukaryotic cells resulting
 CC in cell death. Since the inactive RIP proteins are not cytotoxic to
 CC eukaryotic cells, they can be recombinantly expressed in such cells and
 CC then converted to active RIP proteins.
 SQ Sequence 250 AA;
 W25146 Length: 250 February 11, 2000 15:48 Type: P Check: 7681 ..
 1 ANVSFSLGSA DSKSYSKFIT ALRKALPSKE KVSNIPLLP SASGASRYTL
 51 MOISNDARA ITMAIDVTNV YIMGYLVNST SYFANESDAK LMSQYFKGS
 101 TLVTIPSYGN YERLONAGK IREKIPLGR ALDSALTSIF HYDSTAAAF
 151 FVYLIQTFAE ASREKYEGQ IIERIPKNEV PSPAALSLEN EAWSLSKOI
 201 QLAOTNGAF RTPVYITDKR GQVEITVNTA SKVOIKDVNS KLLNKNONIA
 11AA_SEQUENCE 1.0

ID W22410 standard: Protein; 140 AA.
AC W22410;
DE 08-DEC-1997 (first entry)
KW Alpha-4 integrin mouse Mab 21.6 VH region.
KW asthma; atherosclerosis; AIDS; dementia; diabetes; tumour;
KW metastasis; inflammatory bowel disease; rheumatoid arthritis;
KW transplant rejection; graft versus host disease; nephritis;
KW atopic dermatitis; psoriasis; myocardial ischaemia;
KW acute leukocyte mediated lung injury; therapy.
OS Mus musculus.
FH Key Location/Qualifiers
FT peptide 1..19 /label= leader
FT region 20..49 /label= FR1
FT region 50..54 /note= "framework region 1"
FT region 55..68 /note= "complementarity determining region 1"
FT region 69..85 /label= FR2
FT region 86..117 /note= "complementarity determining region 2"
FT region 118..131 /label= FR3
FT region 132..140 /note= "complementarity determining region 3"
FT region /label= FR4 /note= "framework region 4"
PD WO9718838-A1.
PN 29-MAY-1997.
PF 21-NOV-1996; U18807.
PR 21-NOV-1995; US-561521.
PA (ATHE-) ATHENA NEUROSCIENCES INC.
PI Bendig MM, Jones ST, Leger OJ, Saidanba J, Yednock TA;
DR WPI: 97-297879/27.
DR N-PSDB: T74760.
PT Uses of humanised alpha-4 integrin antibody - for treatment of
PT asthma, atherosclerosis, AIDS dementia, etc.
PS Claim 18: page 69-70/107pp; English.
CC This polypeptide comprises the heavy chain variable region (VH) of
CC mouse anti-alpha-4 integrin monoclonal antibody 21.6. The
CC complementarity determining regions (CDRs) of the 21.6 VH can be
CC incorporated into a human 21/28'CL framework to produce a claimed
CC humanised 21.6 VH (see W22413) and a claimed humanised 21.6
CC antibody that is used in the manufacture of a medicament for
CC treating a disease selected from asthma, atherosclerosis, AIDS,
CC dementia, diabetes, inflammatory bowel disease, rheumatoid
CC arthritis, transplant rejection, graft versus host disease, tumour
CC metastasis, nephritis, atopic dermatitis, psoriasis, myocardial
CC ischaemia, and acute leukocyte mediated lung injury. The antibody
CC may also be used in the affinity purification of alpha-4 integrin
CC for use as a vaccine or an immunogen. It is also useful for
CC generating idiotypic antibodies. The humanised antibodies of the
CC invention have a half-life in the human circulation essentially
CC equivalent to that of naturally occurring human antibodies.
SQ Sequence 140 AA;
W22410 Length: 140 February 11, 2000 15:48 Type: P Check: 2629 ..
1 MGSWMFPL MAVYGVNVE VOLOQSGAEI VEPGASVLS CTAGFPIKD
51 TIHCVKQRP EGGLEMIGRI DPANGTKYD PKFOCKAITT ADISSNTAYL
101 QSSLSLEST AVYFCARCY YGVYGYAMD YWGCTSVTV
11AA_SEQUENCE 1.0

ID W22413 standard: Protein; 123 AA.
AC W22413;
DE 08-DEC-1997 (first entry)
KW Humanised alpha-4 integrin antibody 21.6 VH Ha.
KW Alpha-4 integrin: humanised antibody; monoclonal antibody 21.6;
KW asthma; atherosclerosis; AIDS; dementia; diabetes; tumour;
KW metastasis; inflammatory bowel disease; rheumatoid arthritis;
KW transplant rejection; graft versus host disease; nephritis;
KW atopic dermatitis; psoriasis; myocardial ischaemia;
KW acute leukocyte mediated lung injury; therapy.
OS Chimeric Mus musculus;
OS Chimeric Homo sapiens;
OS Chimeric synthetic.
FH Key Location/Qualifiers
FT region 1..30 /label= FR1
FT region /note= "21/28'CL framework region 1"
FT region 31..35 /note= "21/28'CL residues 27-30 are replaced by
FT region 36..49 /note= "those of Mab 21.6, involved in antigen
FT region 50..66 /label= CDR1 binding"
FT region 67..98 /note= "21.6 complementarity determining region 1"
FT region /label= FR2
FT region 99..112 /note= "21/28'CL framework region 2"
FT region 113..123 /label= CDR3
FT region /note= "21.6 complementarity determining region 2"
FT region /label= FR3
FT region /note= "21/28'CL framework region 3"
FT region 124..125 /note= "21/28'CL Arg-72 is substid. by Ala of mouse
FT region 126..127 /note= "21.6 VL, important in supporting the CDR2
FT region 128..129 loop"
FT region 130..131 /label= CDR3
FT region 132..133 /note= "21.6 complementarity determining region 3"
FT region /label= FR4 /note= "21/28'CL framework region 4"
PD WO9718838-A1.
PN 29-MAY-1997.
PF 21-NOV-1996; U18807.
PR 21-NOV-1995; US-561521.
PA (ATHE-) ATHENA NEUROSCIENCES INC.
PI Bendig MM, Jones ST, Leger OJ, Saidanba J, Yednock TA;
DR WPI: 97-297879/27.
PT Uses of humanised alpha-4 integrin antibody - for treatment of
PT asthma, atherosclerosis, AIDS, dementia, etc.
PS Claim 25: Fig 7; 107pp; English.
CC This polypeptide, designated Ha, comprises the heavy chain variable
CC region (VH) of a humanised alpha-4 integrin monoclonal antibody 21.6. It is
CC composed of complementarity determining regions (CDRs) from the VH
CC region (see W22410) of mouse alpha-4 integrin monoclonal antibody
CC 21.6 and a modified human 21/28'CL framework. It can be expressed
CC in mammalian host cells following PCR amplification and mutagenesis
CC of appropriate fragments of mouse and human DNA sequences. The
CC humanised 21.6 VH and a humanised 21.6 VL (see W22412) can be used
CC to produce a claimed humanised 21.6 antibody that is useful in the
CC manufacture of a medicament for treating asthma, atherosclerosis,
CC AIDS, dementia, diabetes, inflammatory bowel disease, rheumatoid
CC arthritis, transplant rejection, graft versus host disease, tumour
CC metastasis, nephritis, atopic dermatitis, psoriasis, myocardial
CC ischaemia, and acute leukocyte mediated lung injury. The antibody
CC may also be used in the affinity purification of alpha-4 integrin
CC for use as a vaccine or an immunogen. It is also useful for
CC generating idiotypic antibodies. The humanised antibody has a
CC half-life in the human circulation essentially equivalent to that
CC of naturally occurring human antibodies.
SQ Sequence 123 AA;

W22413 Length: 123 February 11, 2000 15:48 Type: P Check: 9557

1 OVOLVOSGAE VKRPGASVSK SCRASEFNK DTYHWRAO PQORLEMKR
51 IDPANGYTRY DKFQGRVTV TADTSASTAY MEISSLASED TAVYICARG
101 XYGNVGVYAN DYWGQGLTVT VSS

!!AA_SEQUENCE 1.0
ID W22428 standard; Protein: 142 AA.

AC W22428; (first entry)
DE Humanised alpha-4 integrin antibody 21.6 VL version Ha.
KW Alpha-4 integrin; humanised antibody; monoclonal antibody 21.6;
KW asthma; atherosclerosis; AIDS; dementia; diabetes; tumour;
KW metastasis; inflammatory bowel disease; rheumatoid arthritis;
KW transplant rejection; graft versus host disease; nephritis;
KW atopic dermatitis; psoriasis; myocardial ischemia;
KW acute leukocyte mediated lung injury; therapy.

OS Chimeric Homo sapiens;
OS Chimeric synthetic.

FR Key location/Qualifiers

FT peptide 1, 19
FT /label= leader
FT 20, 142
FT /label= Mat.protein
FT /note= "VH version Ha (claim 25)"

FT region 20, 49
FT /label= FR1
FT /note= "21/28/CL framework region 1"

FT region 50, 55
FT /label= CDR1
FT /note= "21.6 complementarity determining region 1"

FT region 55, 67
FT /label= FR2
FT /note= "21/28/CL framework region 2"

FT region 68, 85
FT /label= CDR2
FT /note= "21.6 complementarity determining region 2"

FT region 86, 117
FT /label= FR3
FT /note= "21/28/CL framework region 3"

FT region 118, 131
FT /label= CDR3
FT /note= "21.6 complementarity determining region 3"

FT region 132, 142
FT /label= FR4
FT /note= "21/28/CL framework region 4"

FN W09718938-A1.
PD 29-MAY-1997
PF 21-NOV-1996; US-561521
PR (ATHE-) ATHENA NEUROSCIENCES INC.
PA Bendig MM, Jones ST, Leger OJ, Saldanha J, Yednock TA;
PI WPI: 97-287879/27.
DR N-PSDB: 174789.

PT Uses of humanised alpha-4 integrin antibody - for treatment of
PT asthma, atherosclerosis, AIDS, dementia, etc.
PT Example 6; Fig 11, 107pp; English.
CC This polypeptide, designated Ha, comprises the heavy chain variable
CC region (VH) of a humanised alpha-4 integrin antibody 21.6 (see also
CC W22413). It is composed of complementarity determining regions from
CC the VH region (see W22410) of mouse alpha-4 integrin monoclonal
CC antibody 21.6 and a modified human 21/28/CL framework. It can be
CC expressed in mammalian host cells following PCR amplification and
CC mutagenesis of appropriate mouse and human DNA sequences. The
CC humanised 21.6 VH and a humanised 21.6 VL (see W22412) can be used
CC to produce a claimed humanised 21.6 antibody that is useful in the
CC manufacture of a medicament for treating asthma, atherosclerosis,
CC AIDS, dementia, diabetes, inflammatory bowel disease, rheumatoid
CC arthritis, transplant rejection, graft versus host disease, tumour

CC metastasis, nephritis, atopic dermatitis, psoriasis, myocardial
CC ischaemia, and acute leukocyte mediated lung injury. The humanised
CC antibody has a half-life in the human circulation essentially
CC equivalent to that of naturally occurring human antibodies.
SQ Sequence 142 AA;

W22428 Length: 142 February 11, 2000 15:48 Type: P Check: 9019

1 MDWTRVFL LAVAFAHSQ VOLVOSGAEV KKPASVKS CRASGFNMD
51 TYHWRAOP GORLEMMGRI DPANGYTRYD PKFQGRVTVI ADTSASTAYM
101 ELSLSRSEDV AVYICARGY YGNVGVYAND YWGQGLTVTV SS

!!AA_SEQUENCE 1.0
ID W26413 standard; Protein: 121 AA.

AC W26413; (first entry)
DE 16-DEC-1997
DE Swinepox virus HindIII C encoded protein C22R.
KW SPV; vaccine; vector; pseudorabies virus; C22R.
OS Swinepox virus.
PN US5651972-A.

PD 29-JUL-1997
PF 21-APR-1989; 342212.
PR 01-JUL-1992; US-908241.
PR 21-APR-1989; US-342212.
PR 29-JUN-1992; US-908630.
PR 14-SEP-1994; US-307499.
PA (UYFL) UNIV FLORIDA RES FOUND INC.
PI GABbs Epy, Moyer RW, Vinuela E;
DR WPI: 97-392897/36.
DR N-PSDB: T84564.

PT Recombinant swinepox virus vector - used particularly for vaccines
PT against infectious agents, including pseudorabies
PT Example 2; Column 45-48; 70pp; English.
CC This sequence comprises a polypeptide, designated C22R, encoded
CC by the sense strand (open reading frame 3) of the HindIII C
CC fragment (see T84564) of swinepox virus (SPV). C22R shows no
CC apparent homology to known protein sequences. A claimed
CC recombinant vector comprises a heterologous nucleotide sequence
CC inserted into, or replacing, all or a portion of a non-essential
CC SPV gene or nucleic acid sequence of the HindIII C fragment. The
CC vector can be used for the expression of heterologous proteins,
CC both in vivo as a vaccine, and in vitro for production of the
CC selected protein. The heterologous protein is preferably
CC pseudorabies virus gp50 or gp63 for use in swine vaccines. As SPV
CC is host-restricted to swine, the use of modified recombinant SPV as
CC a live vaccine vector eliminates the risk of spreading infection
CC with the virus to other animal populations.
SQ Sequence 121 AA;

W26413 Length: 121 February 11, 2000 15:48 Type: P Check: 3737

1 MIDVITSLD GGCNCKINR HAKSOCKRESS LTISSLENNV PGINPRFIQ
51 NIANEPEKN IPSTANATT LSPYTLILDI HFLAOSAFEE THSIVSIALN
101 KCIFSLSLT YSINKLYIS E

!!AA_SEQUENCE 1.0
ID W17971 standard; Protein: 47 AA.
AC W17971;
DE 12-DEC-1997 (first entry)
DE RAC protein kinase C-terminal binding protein C-terminal region.
KW RAC protein kinase C-terminal binding protein; CRBP; modulator;
KW signal transduction; insulin; cell proliferation; glycogen.
OS Homo sapiens.
PN W09718903-A1.
PD 22-MAY-1997
PF 05-NOV-1996; E04810
PR 15-DEC-1995; GB-025704.
PR 16-NOV-1995; GB-023379.
PA (NOVS) NOVARTIS AG.

PI Hemmlings BA;
DR WPI; 97-289279/26.
DR N-PSDB; 167134.
PT RAC protein kinase, or modulator excluding wortmannin and vanadate,
PT for use as medicament - and screening potential modulators of
PT insulin mediated intracellular signalling using RAC-PK, or fragment
PS Claim 8; Page 25; 38pp; English.
CC This polypeptide comprises the putative C-terminal region of a
CC novel RAC protein kinase C-terminal binding protein (CTBP). Its
CC sequence was deduced from a cDNA clone (767134) isolated from a
CC HeLa library following a yeast two-hybrid screen assay for clones
CC that showed specific interaction with RAC's kinase domain with its
CC C-terminal extension. The C-terminal domain of RAC protein kinase
CC is phosphorylated in response to insulin activation, suggesting a
CC role for CTBP as a modulator of insulin action. RAC protein kinase
CC (see W17972) and modulators of insulin-mediated intracellular
CC signalling such as CTBP can be used in the treatment of
CC abnormalities of cellular metabolism, diseases involving an anomaly
CC in insulin response, and diseases involving an anomaly in glycogen
CC metabolism.
SQ Sequence 47 AA;

W17971 Length: 47 February 11, 2000 15:48 Type: P Check: 6033 ..
1 NSARARASNA PRPGALNSC AARLRTGAL CRPPVGRUP EATRDPS

IIAA_SEQUENCE 1.0
ID W34668; standard; Protein; 438 AA.
AC W34668;
DE 13-FEB-1998 (first entry)
DE Arabidopsis thaliana Rec-A protein.
KW Rec-A protein; Synechococcus recA gene; transit peptide;
KW homologous recombination; DNA repair; the recA gene.
OS Arabidopsis thaliana.
FH Key Location/Qualifiers
FT Peptide 1..50
FT /label= signal-peptide
FT /note= "putative chloroplast transit peptide"
FT /label= mature-peptide
FT Protein 51..438
FT US5674992-A.
FT 07-OCT-1997.
FT 16-JUN-1994; 938332.
FT 28-AUG-1992; US-938332.
PR (CORR.) CORNELL RES FOUND INC.
PI Cerutti H, Jagendorf A;
DR N-PSDB; 193777/46.
DR N-PSDB; 193777
PT DNA encoding chloroplast RecA-like protein from Arabidopsis thaliana
PT - useful for increasing the frequency of homologous recombination
PT recombination when lacking the transit peptide sequence
PS Claim 2; Columns 7-9; 14pp; English.
CC The present sequence is of a Rec-A like protein from Arabidopsis
CC thaliana. The cDNA encoding this protein (193777) was isolated from an
CC Arabidopsis library screened with a Synechococcus recA gene. The cDNA
CC sequence is truncated at its 5' end, and the start of the open reading
CC frame, and thus the start of the transit peptide, is missing. In E. coli,
CC and many other prokaryotes, the Rec-A protein is essential for homologous
CC recombination, and DNA repair. In view of this role, it is likely that
CC the enzyme encoded by the present sequence may also be concerned
CC with DNA repair. The 5' end of the recA gene, i.e. the portion encoding
CC the transit peptide part of the protein that directs the final product
CC into chloroplasts, could be removed and replaced by an amino acid sequence
CC that directs completed proteins to enter the nucleus. By introducing the
CC RecA protein into the nuclei of higher plants, it should be possible to
CC increase the frequency of homologous recombination in the chromosomal
CC genes of the plants.
SQ Sequence 438 AA;

W34668 Length: 438 February 11, 2000 15:48 Type: P Check: 9097 ..
1 DSOLVLSLKL NPSFTPLSPL PFTFCSSFS PSLSFSCIS RLISPYWY

51 AAKLSHKIS SEEDRINGA LSPDASRFL DRKALEAM NDINSSFGK
101 SVTRLSGAG ALVEFFSSGI LTLDLALGG LPRGRVETIY GPSSGRTTL
151 ALHAIKAVOK LGONMAYDA EHAPDPAVK ALGVDENI I VQDPNGEMA
201 LETADRRCRS GAVDLICVDS VSALTPEAEI EGEIGQWKG LQARLMSQAL
251 RKMGSNASKA GCTLIPLNFI RYKIGVYGN PEVTSIGIAL KFPASVRLFI
301 RSARKINSRK GDEDIGLRAR VRVOKSVSR PYKQAEFEIM FGEVSKLGC
351 VLDCAEIMEV VYKGSWSY EDQRLGGRK KALQHLRENP ALQDEIEKKV
401 RLMLDGEVH RSTPLMSSSS SSASHREEE EDSLDOFO

IIAA_SEQUENCE 1.0
ID W24491; standard; peptide; 122 AA.
AC W24491;
DE 27-FEB-1998 (first entry)
DE Novel amylase inhibitor protein subunit isolated from wheat.
KW Amylase inhibitor; wheat flour; wheat; human alpha-amylase;
KW blood sugar level; insulin secretion; fatty acid level; appetite
KW suppressor; treatment; diabetes; prevention; obesity.
OS Triticum sp.
PN EP-785214-A2.
PE 23-JUL-1997.
PE 10-JAN-1997; 100316.
PR 18-JAN-1996; JP-023446.
PA (NAGA-) NAGATA SANGYO CO LTD.
PI Miyazaki T, Morimoto T, Murayama R.
DR WPI; 97-365900/34.
PT New isolated proteins from wheat active as amylase inhibitors - used
PT in the treatment and prevention of diabetes and in the prevention of
PT obesity.
PS Claim 1; Page 16; 22pp; English.
CC This sequence represents a novel amylase inhibitory protein which is
CC composed of two subunits, each having the present sequence. This protein
CC can be prepared by extraction from wheat flour, wheat or wheat gluten.
CC The protein is spherical and has S-S bonds, and has a high inhibitory
CC activity against human alpha-amylase. It has about 5-30 times higher
CC activity as compared with several known amylase inhibitors. It can also
CC inhibit an increase in blood sugar levels, inhibit insulin secretion and
CC feeling of satiety after meals, and easily suppressing appetite. The
CC proteins can be used in the treatment and prevention of diabetes and the
CC prevention of obesity.
SQ Sequence 122 AA;

W24491 Length: 122 February 11, 2000 15:48 Type: P Check: 1423 ..

1 SGPMWCYGY AFVYPALPGC RPYLQLQNG SQVPEALRD CCOQLADISE
51 MCRGALYSM LDMYKHEGV QEGQAGTAF PSCHREYVKL TAASTAVCK
101 LPIVIDASGD GAVYCKGVAA YP

IIAA_SEQUENCE 1.0
ID W24492; standard; peptide; 124 AA.
AC W24492;
DE 27-FEB-1998 (first entry)
DE Novel amylase inhibitor protein 2 subunit isolated from wheat.
KW Amylase inhibitor; wheat flour; wheat; human alpha-amylase;
KW blood sugar level; insulin secretion; fatty acid level; appetite
KW suppressor; treatment; diabetes; prevention; obesity.
OS Triticum sp.
PN EP-785214-A2.
PE 23-JUL-1997.
PE 10-JAN-1997; 100316.
PR 18-JAN-1996; JP-023446.
PA (NAGA-) NAGATA SANGYO CO LTD.
PA (NISS) NISSHIN FLOUR MILLING CO.

PI Miyazaki T, Morimoto T, Murayama R:
 DR WPI: 97-365900/34.
 PT New isolated proteins from wheat active as amylase inhibitors - used
 PT in the treatment and prevention of diabetes and in the prevention of
 PT obesity.
 PS Claim 5; Page 17; 22pp; English.
 CC This sequence represents a novel amylase inhibitory protein which is
 CC composed of two subunits, each having the present sequence. This protein
 CC can be prepared by extraction from wheat flour, wheat or wheat gluten.
 CC The protein is spherical and has S-S bonds, and has a high inhibitory
 CC activity against human alpha-amylase. It has a higher inhibitory
 CC activity as compared with several known amylase inhibitors. It can also
 CC inhibit increases in blood sugar levels, inhibit insulin secretion and
 CC inhibit an increase in fatty acid levels in the blood, thus maintaining a
 CC feeling of satiety after meals, and easily suppressing appetite. The
 CC proteins can be used in the treatment and prevention of diabetes and the
 CC prevention of obesity.
 SQ Sequence 124 AA.

W24492 Length: 124 February 11, 2000 15:48 Type: P Check: 2685 ..

1 SGPWMCYPGY AFKVPALPGC RPYKLCQNG SQVPEAVLRD CCOQLADISE
 51 WCRGALYSM LDSMYKEHGV QEGQAGTGAF PSRCREYVKL TAASTAVCK
 101 LPYVDASGD GAYVCKGVAA YPDA

!!AA_SEQUENCE 1.0
 ID W24493 standard; peptide; 124 AA.
 AC W24493;
 DT 27-FEB-1998 (first entry)
 DE Novel amylase inhibitor protein 3 subunit isolated from wheat.
 KW Amylase inhibitor; wheat flour; wheat; human alpha-amylase;
 KM blood sugar level; insulin secretion; fatty acid level; appetite
 KW suppressor; treatment; diabetes; prevention; obesity.
 OS Trilicium sp.
 PN EP-785214-A2.
 PD 23-JUL-1997.
 PE 10-JAN-1997; 100316.
 PR 18-JAN-1996; JP-023446.
 PA (NAGA-) NAGATA SANGYO CO LTD.
 PI (NISS) NISSHIN FLOUR MILLING CO.
 PI Miyazaki T, Morimoto T, Murayama R:
 DR WPI: 97-365900/34.
 PT New isolated proteins from wheat active as amylase inhibitors - used
 PT in the treatment and prevention of diabetes and in the prevention of
 PT obesity.
 PS Claim 5; Page 18; 22pp; English.
 CC This sequence represents a novel amylase inhibitory protein which is
 CC composed of two subunits, each having the present sequence. This protein
 CC can be prepared by extraction from wheat flour, wheat or wheat gluten.
 CC The protein is spherical and has S-S bonds, and has a high inhibitory
 CC activity against human alpha-amylase. It has about a higher inhibitory
 CC activity as compared with several known amylase inhibitors. It can also
 CC inhibit increases in blood sugar levels, inhibit insulin secretion and
 CC inhibit an increase in fatty acid levels in the blood, thus maintaining a
 CC feeling of satiety after meals, and easily suppressing appetite. The
 CC proteins can be used in the treatment and prevention of diabetes and the
 CC prevention of obesity.
 SQ Sequence 124 AA.

W24493 Length: 124 February 11, 2000 15:48 Type: P Check: 3320 ..

1 SGPWMCYPGY AFKVPALPGC RPYKLCQNG SQVPEAVLRD CCOQLADISE
 51 WCRGALYSM LDSMYKEHGV QEGQAGTGAF PSRCREYVKL TAASTAVCK
 101 LPYVDASGD GAYVCKGVAA YPDA

!!AA_SEQUENCE 1.0
 ID W29523 standard; Protein; 124 AA.
 AC W29523;
 DT 09-MAR-1998 (first entry)

DE Wheat amylase inhibitor 0.26 Ala subunit.
 KW Amylase inhibitor; 0.26 Ala; wheat; visceral fat; obesity; therapy.
 OS Trilicium aestivum.
 PN EP-784978-A2.
 PD 23-JUL-1997.
 PE 11-JAN-1997; 100399.
 PR 18-JAN-1996; JP-023445.
 PA (NAGA-) NAGATA SANGYO CO LTD.
 PI (NISS) NISSHIN FLOUR MILLING CO.
 PI Goda T, Miyazaki T, Morimoto T, Murayama R, Takase S:
 DR WPI: 97-365735/34.
 PT Use of an amylase inhibitor from wheat - for inhibiting the
 PT accumulation of visceral fat to prevent visceral fat obesity
 PS Claim 2; Page 16; 22pp; English.
 CC This protein comprises the subunit of the wheat amylase
 CC inhibitor homodimer protein 0.26 Ala. An agent for inhibiting an
 CC accumulation of visceral fat and hence for preventing visceral
 CC fat obesity is claimed. It contains an amylase inhibitor of wheat
 CC origin as an active ingredient. The amylase inhibitor is selected
 CC from 0.26 Ala, 0.26 Alb (see W29524) and 0.19 Al (see W29525). The
 CC 0.26 Ala protein has inhibitory activity of 26,100 U/mg against
 CC human pancreatic alpha-amylase. Also claimed is a food additive
 CC comprising 0.26 Alb. The amylase inhibitors can inhibit the
 CC digestion of starch, inhibit increases in blood glucose level,
 CC reduce insulin secretion and reduce the activity of lipogenic
 CC enzymes in the viscous.
 SQ Sequence 124 AA.

W29523 Length: 124 February 11, 2000 15:48 Type: P Check: 2685 ..

1 SGPWMCYPGY AFKVPALPGC RPYKLCQNG SQVPEAVLRD CCOQLADISE
 51 WCRGALYSM LDSMYKEHGV QEGQAGTGAF PSRCREYVKL TAASTAVCK
 101 LPYVDASGD GAYVCKGVAA YPDA

!!AA_SEQUENCE 1.0
 ID W29524 standard; Protein; 122 AA.
 AC W29524;
 DT 09-MAR-1998 (first entry)
 DE Wheat amylase inhibitor 0.26 Alb subunit.
 KW Amylase inhibitor; 0.26 Alb; wheat; visceral fat; obesity; therapy.
 OS Trilicium aestivum.
 PN EP-784978-A2.
 PD 23-JUL-1997.
 PE 11-JAN-1997; 100399.
 PR 18-JAN-1996; JP-023445.
 PA (NAGA-) NAGATA SANGYO CO LTD.
 PI (NISS) NISSHIN FLOUR MILLING CO.
 PI Goda T, Miyazaki T, Morimoto T, Murayama R, Takase S:
 DR WPI: 97-365735/34.
 PT Use of an amylase inhibitor from wheat - for inhibiting the
 PT accumulation of visceral fat to prevent visceral fat obesity
 PS Claim 2; Page 17; 22pp; English.
 CC This protein comprises the subunit of the novel wheat amylase
 CC inhibitor homodimer protein 0.26 Alb. An agent for inhibiting an
 CC accumulation of visceral fat and hence for preventing visceral
 CC fat obesity is claimed. It contains an amylase inhibitor of wheat
 CC origin as an active ingredient. The amylase inhibitor is selected
 CC from 0.26 Alb, 0.26 Ala (see W29523) and 0.19 Al (see W29525). The
 CC 0.26 Alb protein has inhibitory activity of 20,500 U/mg against
 CC human pancreatic alpha-amylase. Also claimed is a food additive
 CC comprising 0.26 Alb. The amylase inhibitors can inhibit the
 CC digestion of starch, inhibit increases in blood glucose level,
 CC reduce insulin secretion and reduce the activity of lipogenic
 CC enzymes in the viscous.
 SQ Sequence 122 AA.

W29524 Length: 122 February 11, 2000 15:48 Type: P Check: 1423 ..

1 SGPWMCYPGY AFKVPALPGC RPYKLCQNG SQVPEAVLRD CCOQLADISE
 51 WCRGALYSM LDSMYKEHGV QEGQAGTGAF PSRCREYVKL TAASTAVCK

101 LPIVDASGD GAYVCKGVAA YP

11AA_SEQUENCE 1.0
 ID W29525 standard; Protein; 124 AA.
 AC W29525;
 DT 09-MAR-1998 (first entry)
 DE Wheat amylase inhibitor 0.19 Al subunit.
 KW Amylase inhibitor; 0.19 Al; wheat; visceral fat; obesity; therapy.
 RN Litchium aestivum.
 PS EP-784978-42.
 FD 23-JUL-1997.
 PR 11-JAN-1997; 100399.
 PR 18-JAN-1996; JP-023445.
 PA (NAGA-) NAGATA SANGYO CO LTD.
 PI (NITS) NISSHIN FLOUR MILLING CO.
 DR Goda T, Miyazaki T, Morimoto T, Murayama R, Takase S;
 WPI: 97-365735/34.
 PT Use of an amylase inhibitor from wheat - for inhibiting the
 accumulation of visceral fat to prevent visceral fat obesity
 PS Claim 2; Page 18; 22pp: English.
 CC This protein comprises the subunit of the wheat amylase
 inhibitor homodimer protein 0.19 Al. An agent for inhibiting an
 accumulation of visceral fat and hence for preventing visceral
 fat obesity is claimed. It contains an amylase inhibitor of wheat
 origin as an active ingredient. The amylase inhibitor is selected
 from 0.26 Ala (see W29523), 0.26 Alb (see W29524) and 0.19 Al. The
 0.19 Al protein has inhibitory activity of 20 300 U/mg against
 human pancreatic alpha-amylase. Also claimed is a food additive
 comprising 0.26 Alb. The amylase inhibitors can inhibit the
 digestion of starch, inhibit increases in blood glucose level,
 reduce insulin secretion and reduce the activity of lipogenic
 enzymes in the viscous.
 SC Sequence 124 AA;
 SQ
 W29525 Length: 124 February 11, 2000 15:48 Type: P Check: 3320 ..
 1 SGPMKCYPGQ AFQVPALPAC RPLRLQCG SQVEAVLRD CCOQLAHISE
 51 WCRGALYSM LDMYKEHGA OEGQAGTGF PROCRREVYKL TAASTAVCR
 101 LPIVDASGD GAYVCKGVAA YPDA
 11AA_SEQUENCE 1.0
 ID R63906 standard; Protein; 248 AA.
 AC R63906;
 DT 27-JUL-1995 (first entry)
 DE Type I ribosome-inactivating protein luffin.
 KW Type I ribosome-inactivating proteins; luffin; RIPs;
 KW cytotoxic therapeutic agents; autoimmune disease; cancer;
 KW luffa sp.
 OS Luffa sp.
 PN W09426910-A.
 PD 24-NOV-1994.
 PR 12-MAY-1994; US-064691.
 PR 12-MAY-1993; US-064691.
 PA (XOMA) XOMA CORP.
 FI Better MD, Carroll SS, Studnicka GM, Carroll SF;
 DR WPI: 95-006804/01.
 PT Polynucleotide(s) encoding type I ribosome-inactivating proteins
 PT - which are suitable for use as components of cytotoxic
 PT therapeutic agents.
 PS Disclosures: Fig 4; 22pp: English.
 CC R63906 is the type I ribosome-inactivating protein (RIP)
 luffin, one of the nine RIPs described in R63903-R63911.
 CC RIPs are key components of cytotoxic therapeutic agents (CRAs),
 which include gene fusion products and immunocjugates. CRAs may
 be used to selectively eliminate any cell type to which a RIP
 component is targeted, by the specific binding capacity of the
 second component of the agent. They can be used in the treatment
 of diseases where the elimination of a particular cell type is
 desired, such as autoimmune disease, cancer and graft-versus-host
 disease.

SQ Sequence 248 AA;

R63906 Length: 248 February 11, 2000 15:48 Type: P Check: 7673 ..

1 DVRFSLSSGS STYSKFTIGD LKALPSNGT VYNITILLSS ASGASRYTLM
 51 TLSNDGKAI TVAVDSQLY IMGVLNSTRS YEFNEQAKL ASQYFKGST
 101 IYTLPSGNY EKLQTAGKI REKIPLGFA LDSALTIIFH YDSTAAAF
 151 LVILQITAA SRFYIEGOI IERISNOVP SLATISLENS LMSALSKQIO
 201 LAQTNNGTR TPVITDQK QREVEITNNTS KVTAKNIQL LNKONVA
 11AA_SEQUENCE 1.0
 ID W45093 standard; peptide; 176 AA.
 AC W45093;
 DT 07-MAY-1998 (first entry)
 DE Residues 217-392 of human type B EBNA2 (strain Ag876).
 KW CBFI/RBPJK interaction domain; EBNA2; transplant.
 OS Epstein-Barr virus.
 FH Key
 FT Region
 FT Location/Qualifiers
 FT 69..76
 FT /label= conserved_region_6
 PN W09730081-A1.
 PD 21-AUG-1997.
 PR 12-FEB-1997; 002243.
 PR 13-FEB-1996; US-011557.
 PA (UXXO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
 PI Hayward SD, Ling PD;
 DR WPI: 97-424974/38
 PT Useful with homology to CBFI-RBPJK interaction domain of EBNA2 -
 PT peptide with homology to CBFI-RBPJK interaction domain of EBNA2 -
 PS Disclosures: Fig 5; 62pp: English.
 CC The patent discloses five new peptides which display substantial
 homology with the CBFI/RBPJK interaction domain of EBNA2 (see
 W24280-W24284). In particular, the peptides include amino acids
 318-327 of EBNA2 (i.e. conserved region 6, or C6). The peptides,
 when administered in an amount sufficient to inhibit interaction
 of EBV EBNA2 with CBFI, are used to protect mammals prior to or
 following a transplant, against Epstein-Barr virus. Labeled
 CC peptides are used to detect factors that interact with EBV EBNA2.
 CC The present sequence is from one strain of EB virus (human type B)
 CC used to identify C6 (GPWMPV; residues 69-76).
 SQ Sequence 176 AA;
 W45093 Length: 176 February 11, 2000 15:48 Type: P Check: 9381 ..
 1 HQSLPDRPP PHOSTPHCSS DSGLEPPPT SYSPSMTLS PERLPPAP
 51 AHPLEVIID QOALPPTPGP PMWPPVDRPT PTQTPTPTNT KQPGQGQGR
 101 GRWRGRGRK GRGRMKRLPE PRRPGDISS PSMQSLSPV SLHQGGPPN
 151 SPRPGSTAG PVCRTPTSAT PDISPT
 11AA_SEQUENCE 1.0
 ID W38190 standard; Protein; 107 AA.
 AC W38190;
 DT 11-MAY-1998 (first entry)
 DE Soybean SCARCECROW SRPgl protein.
 KW SCARCECROW; SCR gene; SRPgl; transgenic plant; root; gravitropism;
 KW crop improvement; soybean.
 OS Glycine max.
 FH Key
 FT Location/Qualifiers
 FT 23
 FT Misc.difference /note= "encoded by NNC"
 FT
 PN W09741152-A1.
 PD 06-NOV-1997.
 PR 25-APR-1997; U07022.
 PR 24-APR-1997; US-842445.
 PR 26-APR-1996; US-638617.
 PA (UNIV) UNIV NEW YORK STATE.

PI Benfey PN, DiLaurenzio L, Helariutta Y, Malamy JE,
 PI Pysh L, Mysiocka-Diller J;
 DR WPI: 97-549683/50.
 DR N-PSDB: 195765.
 FT DNA encoding Arabidopsis SCARECROW protein - useful to modify plant
 FT cell division and therefore alter root development, or alter plant
 FT stem or hypocotyl gravitropism
 PS Claim 2; Page 136; 221pp; English.
 CC This is the deduced amino acid sequence of soybean genomic clone
 CC SRP1 (see 195765). It shows homology to Arabidopsis SCARECROW
 CC (SCR) protein (see W38178). SCR is a member of a novel protein
 CC family and contains a number of potential functional domains
 CC similar to those found in transcription factors. SCR is expressed
 CC specifically in embryo root progenitor tissue and in certain root
 CC and stem tissues. It regulates a specific asymmetric division, and
 CC controls gravitropic response in aerial structures and root
 CC formation. SCR proteins (see W38178-201) from dicots and monocots,
 CC SCR proteins lacking 1-4 of motifs I-VI, and polypeptides
 CC corresponding to motif I, II, III, IV, V or VI or SCR are claimed.
 CC Transgenic plants can be engineered to overexpress the SCR protein,
 CC so that cell division is increased in roots, resulting in thicker
 CC root development while a plant with an altered stem or hypocotyl
 CC an antisense molecule that suppresses the expression of endogenous
 CC SCR gene product have thinner root development. A gene of interest
 CC can be placed under control of a SCR promoter and expressed in a
 CC plant to confer herbicide, salt, pathogen or insect resistance, or
 CC when expressed in stems to increase starch, lignin or cellulose
 CC biosynthesis (all claimed).
 SO Sequence 107 AA;
 W38190 Length: 107 February 11, 2000 15:48 Type: P Check: 1840 ..
 1 STENLEDANK MLEISQLSLT PXTYSORVA AYFSEAIR LVSSCLTYA
 51 TLPHHQSHK VASAFQYFNG ISPLVEFSH TANQALQDAF EREBRVHID
 101 LDMOGL
 !!AA_SEQUENCE 1.0
 ID W34214 standard; Protein: 271 AA.
 AC W34214;
 DT 11-MAY-1998 (first entry)
 DE Streptomyces ketoreductase (frenolicin gene P product).
 KM Frenolicin: antibiotic; feed additive; anticoccidial;
 KM coccidiostatic; ketoreductase; Streptomyces roseofulvus.
 OS Streptomyces sp.
 PN EP-806480-A2.
 PD 12-NOV-1997.
 PF 02-MAY-1997; 107329
 PR 07-MAY-1996; US-016753
 PA (HOFF) HOFMANN LA ROCHE & CO AG F.
 PI Reeves CD SOLIDAY CL;
 DR WPI: 97-538619/50.
 DR N-PSDB: 193095.
 FT Streptomyces frenolicin gene cluster - useful for producing
 FT recombinant frenolicin antibiotics
 PS Claim 8; Page 33-34; 66pp; English.
 CC This is a ketoreductase related to the actinorhodin actIII protein.
 CC It is the encoded product of gene P of a claimed Streptomyces
 CC frenolicin gene cluster (see 193095) that contains specifically
 CC claimed coding sequences (genes A-U) that respectively encode 21
 CC proteins (see W34199-219) involved in frenolicin synthesis. The 21
 CC genes of the cluster can be divided into 5 subclusters: (1) genes
 CC A, B, C, D and F encode an efflux pump; (2) genes H, I, J and K
 CC encode butyrate starter synthases; (3) genes L, M and N encode
 CC polyketide synthases; (4) genes O, P, Q and R encode a hemiketalase,
 CC a ketoreductase and cyclases/dehydrases; and (5) genes S and T
 CC encode a keto/enoyl reductase and a hydrolase. Also claimed are
 CC vectors and host cells (especially a Streptomyces sp., particularly
 CC Streptomyces roseofulvus). Host cells can be cultured to produce
 CC frenolicins or frenolicin precursors. The precursors can be
 CC converted to frenolicins by chemical or other methods. The

CC frenolicins can be oxidised to frenolicin B, an antibiotic used as
 CC an anticoccidial agent. The frenolicins can be used as animal feed
 CC additives. 271 AA;
 SO Sequence 271 AA;
 W34214 Length: 271 February 11, 2000 15:48 Type: P Check: 5054 ..
 1 MTTAPHTRP GEACTRGPA LVGATRGIG LAVAALVAR GYPPVVCARD
 51 ABAVARTKE LAAGARVEG VVADYDAAS VHELVAATVA RGPVEVLVN
 101 NAGRSGGGV AELSESLMD VIALNKSYF LVTRVLTG GMGRGVV
 151 NIASTGGKOG VTCAPYAS KRGVVGFTKA LGLELRSGI TVNACPGYV
 201 ETPNAGYVR HYADLMVTE EDVLARFEAK IPIGRYTRPD EVAALVDLYV
 251 TDAAAATQA ALNVCGLGN Y
 !!AA_SEQUENCE 1.0
 ID W41747 standard; protein: 239 AA.
 AC W41747;
 DT 22-MAY-1998 (first entry)
 DE Hepatitis C virus antigen.
 KM non-A non-B hepatitis virus; NANBH; hepatitis C virus; HCV;
 KM antigen; diagnosis; detection.
 OS Hepatitis virus.
 CC PN J05176774.A.
 PD 20-JUL-1993.
 PF 18-DEC-1991; 354708
 PR 18-DEC-1990; JP-412020.
 PA (GREG) GREEN CROSS CORP.
 DR (SHIM/) SHIMOTOHO K.
 DR WPI: 93-260858/33.
 PT Protein contg. non-A non-B hepatitis antigen fragment - prepd. by
 PT culturing transfectants transformed by vector contg. base sequence
 PT coding specified aminoacid sequences, used for detecting hepatitis
 PS Claim 1; Fig 13; 53pp; Japanese.
 CC The present sequence is a non-A non-B hepatitis virus (NANBH) or
 CC hepatitis C virus (HCV) antigen, useful for diagnosis or detection.
 SO Sequence 239 AA;
 W41747 Length: 239 February 11, 2000 15:48 Type: P Check: 3783 ..
 1 VCCSMSTWT GALITPCAAE ESKLPINPLS NSILRHSMV YSTSRASAL
 51 ROKKVFDRDL QVLDHYRDV LKEMKAKAST VKARLSTEE ACKLTPHSA
 101 KSKFGYKAD VASLSRAVN HIRSEWEDL EDTETPIDIT IMAKNEVFCV
 151 QPERGGRKPA RLIVFPDLGV RVCEKMLYD VVSTLPQAVM GPSTGFQYSP
 201 GOREFEVLVN WSKKCPMGF SYDTRCDST VTENDIRTE
 !!AA_SEQUENCE 1.0
 ID W47606 standard; Protein: 194 AA.
 AC W47606;
 DT 11-JUN-1998 (first entry)
 DE HRSV Protein 22k.
 KM HRSV glycoprotein F; gPF; glycoprotein G; gPG; vaccine;
 KM Protein 22k.
 OS human respiratory syncytial virus.
 PN US5716823-A.
 PD 10-FEB-1998.
 PF 12-MAY-1997; 854783.
 PR 13-JUL-1988; US-218737.
 PR 14-JUN-1986; US-818740.
 PR 23-DEC-1986; WO-002756.
 PR 11-JUN-1992; US-897171.
 PR 12-MAY-1997; US-854783.
 PA (PHAA) PHARMACIA & UPJOHN CO.
 PI Collins PL, Wertz GW;
 DR WPI: 98-144802/13.

DR N-PSDB: V18737.
 PT Production of human respiratory syncytial virus glyco-protein F or G
 PT - by culturing eukaryotic host cells transfected with corresponding
 PT DNA
 PS Example 1: Columns 27-30: 17pp; English.
 CC The present sequence was used in the development of a novel method
 CC for the production of human respiratory syncytial virus (HRSV)
 CC glycoprotein F (GPF) or glycoprotein G (GPG). The method comprises
 CC culturing eukaryotic host cells transfected with an isolated DNA
 CC sequence encoding HRSV gpf or gpg. The gp can be used to prepare
 CC vaccine against HRSV.
 SO Sequence 194 AA;

W47606 Length: 194 February 11, 2000 15:48 Type: P Check: 9828 ..

1 MSRRNCKEFG IRGCHLNCR CHFSHNFEM PHALLVARN FMLNRLIKSM
 51 DKSIDLSEI SGAEALDRE EYALGVGV EYISINNI TKOSACYAMS
 101 KLTLEINSD IKLLDNEL NSPKRYNT VASTYESRK NKNQTHLLK
 151 RLPADVLAKT IKNTLDIKS ITINPKESI VSDINDAKN NDTT

!!AA-SEQUENCE 1.0
 ID W50006 standard; Protein: 365 AA.

AC W50006;
 DE 11-JUN-1998 (first entry)
 KW Human papillomavirus-18 E2 protein.
 KW wart; epidermodysplasia verruciformis; laryngeal papilloma;
 KW cervical carcinoma.
 OS Human papillomavirus.
 FT Key
 FT Location/Qualifiers
 FT 1..203
 FT /note="transcription activating domain"
 FT 282..365
 FT /note="DNA-binding domain"
 FT Domain
 FT WO9801148-A1.
 PD 15-JAN-1998.
 PF 08-JUL-1997: U11815.
 PR 09-JUL-1996: US-677206.
 PA (HARD) HARVARD COLLEGE.
 PI Domanick JJ, Howley PM;
 DR WPI: 98-100815/09.
 DR N-PSDB: V18818.
 PT Treating papilloma virus infections, by repressing E6 and E7
 PT oncogenic proteins - using E2 polypeptide or nucleic acid encoding
 PT it, useful for treating warts and cervical cancer
 PS Claim 22: Pages 51-53; 71pp; English.
 CC The present sequence is the human papillomavirus-18 (HPV-18)
 CC E2 protein.
 CC Papillomavirus infections can be treated by administering an E2
 CC protein, comprising a DNA-binding and transcription activating
 CC domain, or the gene encoding the E2 protein to infected cells. The
 CC method is applicable to livestock, zoo animals, pets and humans,
 CC specifically for the treatment, prevention or reversal of warts,
 CC e.g. plantar, common, Butcher's common, flat or genital warts, or
 CC epidermodysplasia verruciformis, or growth of papillomavirus
 CC transformed cells, especially laryngeal papilloma or cervical
 CC carcinoma.
 SO Sequence 365 AA;

W50006 Length: 365 February 11, 2000 15:48 Type: P Check: 5567 ..

1 MOPKRETLSE RLSGVODKII DRYENDSKDI DSOIQWQLI RMENALFFAA
 51 REHGIGTLNH QVPAYNISK SKAHKALELO MALOGIAOSR YKREMDTLID
 101 JOELMNTER THCFKKGOT VOYYPGONKD NCMATYVADS VYIMTDAGTW
 151 DKATCVSHR GLYYKESGN TFYIEKSEC EKYSGTGWIE VHGNNVIOC
 201 NDSMCTSDO TVSATOLYKO LOHTPSPYSS TVSGTAKTY GQISAATRPB

251 HCGIAERQHC GPVNPILGAA TPTGNKRRK LSGNTTPII HUKODRSLK
 301 CLRRLRKHS DMYRDISSTM HMTGANGKNT GILTYVHSE TORRFLNTV
 351 AIPDSVOILV GYWTM

!!AA-SEQUENCE 1.0
 ID W55590 standard; Protein: 307 AA.

AC W55590;
 DE 07-JUL-1998 (first entry)
 KW H. pylori ORF 09cpi0224.429510.c2.46.a cell envelope OMP.
 KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
 KW identification; binding compound; bacteria; life cycle; activator;
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
 OS Helicobacter pylori.
 PN WO9737044-A1.
 PD 09-OCT-1997.
 PP 27-MAR-1997: U05223.
 PR 06-DEC-1996: US-761318.
 PR 29-MAR-1996: US-625811.
 PR 02-APR-1996: US-738731.
 PR 23-OCT-1996: US-738905.
 PR 28-OCT-1996: US-738859.
 PA (ASTRA) ASTRA AB.
 PI Alm RA, Smith DJ;
 DR WPI: 97-503122/46.
 DR N-PSDB: V24999.
 PT Helicobacter pylori nucleic acid sequences and encoded
 PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori
 PT infection and for diagnosis of H. pylori infection
 PS Claims 14,80; Pages 792-793; 1145pp; English.
 CC This sequence is a H. pylori cell envelope outer membrane protein (OMP)
 CC having a C-terminal tyrosine cluster motif. The protein may be used in a
 CC vaccine to prevent or treat H. pylori infection or to identify H. pylori
 CC polypeptide binding compounds, useful as potential H. pylori life cycle
 CC activators or inhibitors. The DNA and probes derived from it may be
 CC used for the identification of H. pylori in a sample and the diagnosis
 CC of H. pylori infection. Nucleic acid sequences complementary to the DNA
 CC act as antisense sequences and can be used to prevent the translation of
 CC H. pylori mRNA. Antibodies against the abundance and distribution of
 CC immunosays to evaluate the abundance and distribution of
 CC H. pylori-specific antigens. The genomic sequence of H. pylori
 CC (ATCC 55679) was determined from overlapping contigs generated by
 CC mechanically shearing the bacterial DNA. The sequences were analysed
 CC for ORF of at least 180 nucleotides, and the predicted coding regions
 CC defined by computer evaluation. No identify likely H. pylori antigens for
 CC vaccine development, the amino acid sequences predicted from various ORF
 CC were analysed for significant homology to other known or reported
 CC membrane proteins. Having identified and determined the sequences of
 CC interest, particular regions can be isolated from H. pylori by PCR
 CC amplification for recombinant polypeptide production, e.g. in E. coli
 CC hosts.
 SO Sequence 307 AA;

W55590 Length: 307 February 11, 2000 15:48 Type: P Check: 428 ..

1 MKKVLLTLLS LSLSEFLNAE RNFYGLINF AGSYIOGOG SIGKSAEN
 51 ALNOAINNAK NSLPNTKAI RDYONALNAV KDSNKIANRF AGNGSGGIF
 101 NELSIGRYKF LGKKRIGFR HSLFFGYQLG GVSVPAGSL IAPLPYFNT
 151 DLINMTNDR RASQEVERR VKGSLIPYKD MNGRTLDADT LKRASRIHR
 201 KSSGLVIGME LGASTWFSAN NLTFPNQVKS RIFPOLGKRF GVRSSDEYD
 251 IDRYGDENYL GSSSVELGVK VPAFKVNTS DDYGRKLDYK RVSVYLYNT
 301 YNEKNKH

!!AA-SEQUENCE 1.0
 ID W55815 standard; Protein: 271 AA.

AC W55815; 15-JUL-1998 (first entry)
 DE Streptomyces roseofulvus; frenolicin gene cluster protein 16.
 KW Streptomyces roseofulvus; frenolicin gene cluster; frenolicin B;
 antibiotic.
 OS Streptomyces roseofulvus.
 PN J10094395-A.
 PD 14-APR-1998.
 PE 07-MAY-1997; 116652.
 PR 04-APR-1997; US-042933.
 RA (HOFF) HOFMANN LA ROCHE & CO AG F.
 DR WPI: 98-279231/25.
 DR N-PSDB: V25925.
 PR Frenolicin gene cluster - useful as an antibiotic
 PS Claim 8; Page 30; 50pp; Japanese
 CC The present sequence represents a protein from the frenolicin gene
 CC cluster from Streptomyces roseofulvus. The present invention describes:
 CC (1) a vector containing the frenolicin gene cluster DNA sequence
 CC operably connected to an expression control sequence; (2) a host cell
 CC transformed by the above vector; (3) a protein coded by the above DNA
 CC sequence; (4) a method for the preparation of frenolicin or a
 CC biosynthetic intermediate for it in which the above cell is cultured
 CC and frenolicin or its biosynthetic intermediate is isolated from the
 CC culture or the cell; (5) a method for the preparation of frenolicin B
 CC by oxidizing frenolicin; and (6) a method for the preparation of a
 CC feed composition by mixing frenolicin with other components.
 CC Frenolicin B is useful as an antibiotic.
 SQ Sequence 271 AA;
 W55815 Length: 271 February 11, 2000 15:48 Type: P Check: 5054 ..
 1 MTRAPTRP GEAGTRGPA LVGTGRGIG LAVALVALVR GYPVVCARD
 51 AEAARATYKE LAAGARVGEA VVADVTDAAS VHELVATVVA RFGPEVLVN
 101 NAGSGGGVT AELSESLMD VYATNKSVE LVTRVLTGG GMTGRGRCV
 151 NIATGKGQGV VVGAPYSAS KHGVGFTRKA LGEELARSGI TVNACPGYV
 201 ETPTAAQVRR HYADLMVTE EDVLAFFAK IPGRTTRPD EYALVVDLV
 251 TDAANAATQ ALNVCGLGN Y
 11AA SEQUENCE 1.0
 ID W56137 standard; Peptide; 186 AA.
 AC W56137;
 DE 14-JUL-1998 (first entry)
 DE Open reading frame 2 peptide of bacteriophage phi-CPG1.
 KW Circular; bacteriophage phi-CPG1; avian C. psittaci; Chp1;
 KW Chlamydia psittaci strain Guinea pig Inclusion Conjunction; GPIC;
 KW blindness; ocular chronic disease trachoma; vaccine.
 OS Bacteriophage phi-CPG1.
 PD US5741697-A.
 PN 21-APR-1998.
 PE 30-NOV-1995; 565386.
 PR 30-NOV-1995; US-565386.
 PA (UYRP) UNIV ROCHESTER.
 PI Bayoll PM, Sala R;
 DR WPI: 98-260507/23.
 DR N-PSDB: V16865.
 PR Bacteriophage phi-CPG1 genomic DNA - is useful for transforming
 PT mammalian Chlamydia strains
 PS Claim 12; Columns 23-24; 28pp; English.
 CC The present sequence represents the product of the second open reading
 CC frame of a circular bacteriophage, designated phi-CPG1. This
 CC bacteriophage is able to infect Chlamydia psittaci strain Guinea pig
 CC Inclusion Conjunction (GPIC). phi-CPG1 was identified by homology to
 CC a previously characterized Chlamydia bacteriophage of avian C. psittaci
 CC named Chp1. Chlamydia strains can be responsible for blindness through
 CC ocular chronic disease trachoma. phi-CPG1 can be used as a cloning
 CC vector especially as a means to transfer mutant or foreign genes into
 CC Chlamydia. phi-CPG1 can also be used to develop Chlamydia vaccines;

CC diagnostic tests and the study for Chlamydia infection and therapeutic
 CC reagents and strategies in infections.
 SQ Sequence 186 AA;
 W56137 Length: 186 February 11, 2000 15:48 Type: P Check: 7408 ..
 1 NMFQIINTL GSAVGAAG LSEFLGASG VILGYLAQKQ NATAKQARE
 51 QVAFQERMSN TAYORAMEDM KKAGLPMYLA FSKGASGSPA GASWSPNNPV
 101 ESANMSGIAY QRLTYERKM QAEIOLNREQ NRIIRNOAIR EGYLAERDKY
 151 KRVAGVPVAT EMDRTSGIL SSARAFKRL FSRKGR
 11AA SEQUENCE 1.0
 ID W51011 standard; protein; 277 AA.
 AC W51011;
 DE 10-AUG-1998 (first entry)
 DE Human liver carbonyl reductase.
 KW Human carbonyl reductase; HCRD; inflammatory disorder; cancer;
 KW immunological disorder; prostaglandin E.
 OS Homo sapiens.
 PN U5356293-A.
 PD 26-MAR-1998.
 PE 09-DEC-1996; 762129.
 PR 09-DEC-1996; US-762129.
 PA (INCY) INCYTE PHARM INC.
 PI Goll SK, Hillman JL;
 DR WPI: 98-321527/28.
 PR DNA encoding human carbonyl reductase polypeptide - useful for
 PT recombinant production of the enzyme and anti-sense treatment of
 PT diseases associated with the enzyme
 PS Disclosure; Figure 2; 27pp; English.
 CC The invention relates to DNA encoding a human carbonyl reductase protein
 CC (HCRD) sequence. Also claimed are: a probe that hybridizes to the DNA
 CC sequence; an expression vector containing the DNA sequence; and a host
 CC cell containing the vector. The enzyme catalyzes the reduction of
 CC prostaglandin or hydrocortisone, as well as in drug metabolism. The cell
 CC can be used to produce a recombinant human carbonyl reductase protein,
 CC which can be used to treat inflammatory or immunological disorders
 CC resulting from excessive prostaglandin E production, or to prevent cancer
 CC development after exposure to carcinogens. The vector can be used for the
 CC same purpose. The probe can be used to detect carbonyl reductase nucleic
 CC acids in biological samples. Antisense constructs based on the
 CC polynucleotides can be used to control expression of the gene. The
 CC present sequence represents human liver carbonyl reductase.
 SQ Sequence 277 AA;
 W51011 Length: 277 February 11, 2000 15:48 Type: P Check: 7216 ..
 1 MSSGIHALV TGNKKGIGLA IYRDLRFLS GDVYLTRADV TRGOAAVQOL
 51 QAEGISPREH QLDIDLOST RALRDLRKE YGGLDVLVN AGIARVADP
 101 TPEHIOAEVT MKNFEGTRD VOTELLPLIK PQGRVAVSS INSVRLNSC
 151 SPELOQKRS ETTIEELVG LMKKEVEDRK KGVHQKGGP SSAYGVTKIG
 201 VYLSRIHAR KLSQKRGDK ILLNCCPGW VRTDMAGPKA TKSPEGRET
 251 PYLALLPPD AEGPHQFVS EKRVEQW
 11AA SEQUENCE 1.0
 ID W28052 standard; Protein; 269 AA.
 AC W28052;
 DE 27-AUG-1998 (first entry)
 DE Amino acid sequence of phospho-beta-glucosidase (bg1A).
 KW Staphylococcus aureus protein; ribozyme; antisense sequence; control;
 KW Staphylococcal gene; regulatory element; bacterial gene expression;
 KW vaccine; Staphylococcal infection; food poisoning; scaled skin syndrome;
 KW toxic shock syndrome.
 OS Staphylococcus aureus.

Key Location/Qualifiers
 MISC_difference 1..269
 residues designated X are not defined in the specification"

W09730070-A1.
 21-AUG-1997.
 19-FEB-1997; U02318.
 20-FEB-1996; US-011888.
 (SMK) SMITHKLINE BEECHAM CORP.
 Black MT, Burnham MK, Hodgson JE, Knowles DJC, Nicholas RO, Pritt JM, Reichard RW, Rosenberg M, Ward JM.
 WPI: 97-424969/39.
 DR N-PSDB: T84002.
 PT Novel polypeptide(s) from *Staphylococcus aureus* strain WCH29 - used to isolate antimicrobial compounds, and in vaccines against *S. aureus* infection.
 PT Claim 6; Pages 433-434; 989pp; English.
 CC The present sequence represents a *Staphylococcus aureus* protein, that, based on homology with a *Bacillus subtilis* protein, is believed to be a phospho-beta-glucosidase (pgla). The DNA sequence was isolated from a library of clones of *S. aureus* WCH 29 in *Escherichia coli*. The DNA sequence can be used in the construction of ribozymes and antisense sequences to control the expression of *Staphylococcus* genes. The DNA sequence is also useful as a source of regulatory elements for the control of bacterial gene expression. The present protein may be used to produce vaccines to enable a host to produce specific antibodies with antibacterial action. These vaccines and antibodies would protect a host against invasion by *S. aureus*, and conditions relating to *Staphylococcus* infection, e.g. *Staphylococcus* food poisoning, scaled skin syndrome, and toxic shock syndrome.
 CC Sequence 269 AA;
 SO

W28052 Length: 269 February 11, 2000 15:48 Type: P Check: 2273

1 MGLNCLRTSI AMTRIFPND EDVPEBEGIA FYDRIFDELI AQGIEPVYTL
 51 SHEPMPLHA KHYGGRNRE VVDYVHFAR VFERKYDKY TYWMFNEIN
 101 NQMDTNPRIE LMTNSGVALT ENDNEEVLY QVAHHELLAS ALAVALGKEI
 151 NPKRICTMI SHVPIPYSC HPKDMEMQI ANLRFFEPD VQVREYYSY
 201 AKKALARKGY DVGWQGDSD ILQGTXYI GFSTYKSTAN KIDVDTYXEN
 251 NIYNGGCGFI CGESAVRNE

11AA SEQUENCE 1.0
 ID W37816 standard; Protein: 317 AA.
 AC W37816.
 DT 28-AUG-1998 (first entry)
 DE Human secreted apoptosis-related protein hSARP3.
 KW Secreted apoptosis-related protein; SARP; hSARP3; human; prostate cancer; breast cancer; diagnosis; gene therapy.
 OS Homo sapiens.
 FH Key
 FT Peptide
 FT Protein
 FT /label= Sig_peptide
 FT /label= Mat_protein
 PN W09813493-A2.
 PD 02-APR-1998.
 PF 24-SEP-1997; U17154.
 PR 11-OCT-1996; US-028363.
 PR 24-SEP-1996; US-026603.
 PA (LXRB-) LXR BIOTECHNOLOGY INC.
 PI Melkonian H, Umansky S;
 PI WPI: 98-230704/20.
 DR N-PSDB: V19114.
 PT New secreted apoptosis-related proteins - useful for modulating apoptosis, particularly for treatment of prostatic or breast cancer, also for diagnosis and monitoring of disease
 PT Claim 1; Page 56-58; 101pp; English.
 CC This polypeptide comprises human secreted apoptosis-related

protein hSARP3 that modulates apoptosis through cell-cell or cell-extracellular matrix signalling. Its amino acid sequence was deduced from a clone (see V1914) obtained from a human pancreas cDNA library. Murine hSARP1 (see W37814), human hSARP1 (see W37816) and hSARP2 (see W37817) proteins are also claimed. hSARP3 is expressed predominantly in pancreas. SARP polypeptides can be obtained from recombinant host cells. Antibodies specific for SARP polypeptides can be used in immunoassays for detecting levels of expression of SARP, particularly for diagnosis or monitoring of diseases associated with SARP expression. Specifically, they are used to detect cancer of the prostate or breast (by detecting hSARP1 and 2, respectively). SARP polypeptides and nucleic acids can also be used to treat these cancers, or more generally apoptosis-related disease (e.g. infection with HIV or reperfusion injury), also (not claimed) to prevent apoptosis in cultured cells, to improve preservation of organs for transplantation, for in situ preservation for by-pass operations and to treat dermatological disorders. SARP polypeptides can also be used to identify agents, potentially useful therapeutically, that modulate the effects of SARP on mit-1-tyrosine protein interaction.
 CC Sequence 317 AA;
 SO

W37816 Length: 317 February 11, 2000 15:48 Type: P Check: 3813

1 MRAAAGVY RTAALLLG ALHMAPRCE EYDYGWQAE PLHGYSYSP
 51 PQCIDPADL PLCHTVGYR MRLPYLEHE SLAEVYQAS SWPLIARC
 101 HSDQVFLCS LFPVCLDRP IYPCSLCEA VRAAGCAPLME AYGFWPEML
 151 HCHKEPLDND LCIAVQFGL PATAPVYTKI CAQCEHNSA DGLMFQCCSS
 201 DFVAKMIKE IKRNGDRKL IGAOKKKLL KPPLKRRDT KRLVIMKNG
 251 AGCCPQQLDS LAGSFYLMGR KYDQQLLMA VYRWKKRKE MKFAVKEFMS
 301 YPCSLYPTFF YGAAEPH

11AA SEQUENCE 1.0
 ID W39024 standard; Protein: 71 AA.
 AC W39024.
 DT 14-SEP-1998 (first entry)
 DE Enterococin-900.
 KW Enterococin-900; bacteriocin; signal peptide; protein secretion; preservative; food spoilage; lactic acid bacterium; antibacterial.
 OS Enterococcus faecium 900.
 FH Key
 FT Peptide
 FT Protein
 FT /label= Sig_peptide
 FT /label= Mat_protein
 PN W09809639-A1.
 PD 12-MAR-1998.
 PF 05-SEP-1997; U15609.
 PF 05-SEP-1996; US-026257.
 PA (UYAL-) UNIV ALBERTA.
 PI Franz CM, Greer GG, Leisner JJ, McCormick JK, McMullan LM, Pkns J, Poona, Stiles ME, Van Belkum MJ, Vederas JC,
 PI Morobo RJ, Morobo RW;
 DR WPI: 98-193319/17.
 DR N-PSDB: V11704-05.
 PT Bacterial growth inhibiting peptide(s) enterococin 900 or brochocin C - used to inhibit growth of susceptible bacteria in e.g. foodstuff, living animal, food preparation area or fermentation vessel
 PT Example 7; Page 138-139; 174pp; English.
 CC This polypeptide comprises the enterococin-900 pre-protein of Enterococcus faecium 900. Enterococin-900 is a novel broad spectrum bacteriocin that has activity against other strains of Enterococcus as well as many other organisms. The invention includes a method for inhibiting pathogenic bacteria by providing a bacteriocin selected from enterococin-900 and brochocin-C (see W59021), either as a composition or by providing a bacterial source of the bacteriocins. This is used to inhibit spoilage

bacteria in foodstuffs, such as meat, inhibit pathogenic bacteria
topically in animals, e.g. to treat mastitis (claimed), and inhibit
bacterial infection of fermentation reactors.
Sequence 71 AA;

W59024 Length: 71 February 11, 2000 15:48 Type: P Check: 9911 ..

1 MONKELSTR EKKOITIGEN DHRPNEPLNT PNNLSKGAK CGAIAAGLF
51 GIPKPLAMA AALANYSKC N

11AA_SEQUENCE 1.0
ID W57330 standard; Protein: 339 AA.

AC W57330;

DT 14-SEP-1998 (first entry)

DE Glycerol-3-phosphate dehydrogenase gpsa.

KW Glycerol-3-phosphate dehydrogenase; g3pdh; gpsa; yeast.

OS Saccharomyces sp.

PN W09821340-A1.

PD 22-MAY-1998.

PE 10-NOV-1997; U20293.

PR 13-NOV-1996; US-030602.

PA (DUPO) DU PONT DE NEMOURS & CO E I.

PI Butthuis BA, Gatenby AA, Haynie SL, Hsu AK, Lareau RD;

DR WPI: 98-297943/26.

PT Fermentative production of glycerol using recombinant host -

PT containing genes for glycerol-3-phosphate dehydrogenase and/or

PT glycerol-3-phosphatase

PS Claim 9, Page 36-37; 57pp; English.

CC This claimed saccharomycetes polypeptide comprises a

CC glycerol-3-phosphate dehydrogenase (g3pdh) that catalyses the

CC conversion of dihydroxyacetone phosphate to glycerol-3-phosphate.

CC It is encoded by the g3pdh gene. The invention provides recombinant

CC organisms that express g3pdh and/or glycerol-3-phosphatase (g3p)

CC (see also W57324-32) useful for the production of glycerol from a

CC variety of C-sources. A host cell is preferably transformed with a

CC cassette containing either a g3pdh gene and/or a g3p gene and then

CC cultured in the presence of a mono-, oligo-, polysaccharide or

CC IC-substrate. The glycerol obtained is used in cosmetics, liquid

CC soaps, pharmaceuticals, lubricants and antifreezes; its esters are

CC used in the oil and fat industries. The method produces glycerol

CC rapidly and inexpensively without generation of polluting

CC by-products.

SO Sequence 339 AA;

W57330 Length: 339 February 11, 2000 15:48 Type: P Check: 4443 ..

1 MNQNASMTV IGAGSYGTAL AITLARGHE VVLMGDPDH IATLEDRCN

51 AAFLEDPFP DTLHLESDLA TALASRNIL VVPSHVEGE VLROIKPLMR

101 PDARLWATK GLEAETGRLL QDVAREALGD QIPLAYISGP TFAKELAAGL

151 PTAISLASTD QTFADLLOOL LHCKGSPFVY SNDFIGVOL GGAANYVTAI

201 GAGNSDGI GF GANARTALIT RGLAEMSLG AALGADPATF MGAMGLGLV

251 LTCIDNOSRN RRFQMLGGG MDVOSAOKRI GOVVEGYRNT KEVELAHRE

301 GVEMPITEEI YOVLGCKNA REAALLTLGR ARKDESSH

11AA_SEQUENCE 1.0
ID W62641 standard; Protein: 137 AA.

AC W62641;

DT 18-SEP-1998 (first entry)

DE Flea serine protease inhibitor variable domain protein pfspl14-137.

KW Serine protease inhibitor; SPI; anti-haematophagous ectoparasite;

KV vaccine; control; flea; acarina; immune response.

OS Ctenocephalides felis.

PN W09820034-A2.

PD 14-MAY-1998.

PF 03-NOV-1997; U20678.

07-NOV-1996; US-745995.

PA (HESK-) HESKA CORP.

PI Brandt KS, Maddux JD, Silver GM, Wisniewski N;

DR WPI: 98-286864/25.

DR N-PSDB: V38773.

PT New nucleic acid encoding new serine protease inhibitors from fleas

PT - and related antibodies, inhibitors, recombinant viruses and cells,

PT useful for control of fleas on animals, especially cats and dogs

PS Claim 4; Page 151; 176pp; English.

CC The present sequence, designated pfspl14-137, represents Ctenocephalides

CC felis serine protease inhibitor (SPI) variable domain protein. The SPI

CC protein represents a novel target for anti-haematophagous ectoparasitic

CC vaccines and drugs. Compositions containing the inhibitor are used to

CC control fleas (adults, larvae) on mammals or birds, particularly cats

CC and dogs. More generally the composition is also used for control of

CC insects and acarina, on animals or on stored goods, plants, trees etc.

CC including domestic pests and insects that are vectors of disease. The SPI

CC proteins are able to induce a protective immune response against SPI.

SO Sequence 137 AA;

W62641 Length: 137 February 11, 2000 15:48 Type: P Check: 7060 ..

1 EKLQNDLQ NLQRYSYVE VILDPKRI ESEININDPL KKLQSDMEV

51 PKADFKGL EGSDEMLYIS KVIOKAFIEV NEBGAEMAAA TAVFATREVI

101 KYIAKEIFNC DHPFYFALVH SQEGTSAPLF TGAFFTP

11AA_SEQUENCE 1.0
ID W55052 standard; Protein: 118 AA.

AC W55052;

DT 28-SEP-1998 (first entry)

DE Sunflower antifungal protein M559 fragment.

KW Antifungal; fungicide; M559; sunflower; carbohydrate oxidase;

KM glucose oxidase; transgenic plant; Phytophthora; Pythium;

KX crop protection; disease resistance.

OS Helianthus annuus cv. Zedlitz.

FS Key Location/Qualifiers

FT Misc-difference 68 /label= Arg, Lys

PD W09813478-A2.

PN 02-APR-1998.

PF 04-SEP-1997; E04923.

PR 19-MAR-1997; EP-200831.

PR 04-SEP-1996; EP-202466.

PA (MOGE-) MOGEN INT NV.

PI Custers JHHV, Lageweg W, Melchers LS, Ponstein AS,

PI Sela-Buurlage MB, Stulver MH, Van Deventer-Troost JPE;

DR WPI: 98-230692/20.

DR N-PSDB: V27251.

PT New plant proteins having anti-fungal activity - useful as, e.g.

PT carbohydrate oxidase(s) for protection against Phytophthora and

PT Pythium sp.

PS Claim 5; Page 53-54; 139pp; English.

CC This polypeptide is the deduced product of a DNA sequence (see

CC V27251) amplified from sunflower cv. Zedlitz genomic DNA using

CC PCR primers (see V27249-50) based on tryptic peptides (see

CC W53050-51) of the isolated sunflower M59 antifungal protein.

CC A Southern blot of sunflower genomic DNA using this DNA fragment

CC revealed the existence of multiple homologous sequences in the

CC sunflower genome. A full-length clone (see V27260) for sunflower

CC M59 (see W55053) was subsequently obtained. Claimed antifungal

CC proteins, which may include this M59 polypeptide sequence, have

CC a mol.wt. of 55-65 kDa (SDS-PAGE), have carbohydrate oxidase

CC (especially glucose oxidase) activity, show anti-Phytophthora

CC and/or anti-Pythium activity, can be expressed in transgenic

CC plants to reduce susceptibility to infection by fungi, or

CC expressed in host cells for use in antifungal compositions. Plants

CC engineered to express the antifungal proteins require reduced

CC treatments with fungicides and have a longer shelf life.

SO Sequence 118 AA;

W55052 Length: 118 February 11, 2000 15:48 Type: P Check: 4306 ..

1 DPFPPITGEV YTPGNSSFPV VLQNYRNLR ENETTPKPF LITTAHVSH
 51 IOAAVVCQKQ NRLLKTXSG GHDEGLSYL TTNQPPFIY DMENLSINV
 101 DIEQETAMVQ AGATLGEV

11AA SEQUENCE 1.0
 ID W48722 standard; Protein: 355 AA.
 AC W48722:
 DT 25-SEP-1998 (first entry)
 DE Human V28 seven transmembrane receptor.
 KW V28; Placenta; seven transmembrane receptor; 7TM; signal transduction;
 KM Immunology; Inflammation.
 OS Homo sapiens.
 FH Key
 FT Domain Location/Qualifiers
 FT Domain 26..56 "Transmembrane domain 1"
 FT Domain 68..92 "Transmembrane domain 2"
 FT Domain 107..125 "Transmembrane domain 3"
 FT Domain 146..167 "Transmembrane domain 4"
 FT Domain 197..219 "Transmembrane domain 5"
 FT Domain 232..253 "Transmembrane domain 6"
 FT Domain 273..297 "Transmembrane domain 7"
 FT Domain /note= "Transmembrane domain 7"
 PN US5759804-A.
 PD 02-JUN-1998.
 PE 17-NOV-1993: 153848.
 PR 17-NOV-1992: US-977452.
 PA (TCOS-) TCOS CORP.
 PI Godiska R, Gray PM, Schwelckart VL;
 DR N-PSDB: V18343; V18352.
 DT WPI: 98-33232/23.
 PT DNA encoding V28 seven transmembrane receptor polypeptide - useful
 PT for producing recombinant polypeptide and anti-V28 antibodies, and
 PT in screening assays for V28 agonists and antagonists
 PS Claim 1: Columns 61-64; 56pp; English.
 CC The present sequence claimed by the invention represents the human
 CC V28 seven transmembrane (7TM) receptor. It is encoded by the full
 CC length V28 genomic DNA (V18343) or by the V28 cDNA (V18352). As
 CC the V28 protein (W48722) is a 7TM receptor it is probably involved in
 CC signal transduction. The invention also claims that cells transformed
 CC with V28 DNA can be used to produce the recombinant polypeptide, to
 CC produce anti-V28 antibodies or in screening assays for V28 agonists or
 CC antagonists. The antibodies, agonists and antagonists could then be
 CC used to modulate V28 receptor-ligand binding, for e.g. in immunological
 CC and/or inflammatory events in vivo.
 SQ Sequence 355 AA;
 W48722 Length: 355 February 11, 2000 15:48 Type: P Check: 2863 ..

1 MDQPESSATE NERYDLAER CTIGIYVFG TVEISITYSV IFAIGLVNML
 51 LVFVALTNSK KPRSVTDIYI LNLALSDILE VAPLPFTHY LINEKGLVNA
 101 MCKETTAFF IGFPGSIFPI TVISIDRYLA IVLANSMNN RTVQHCVTIS
 151 LGVMAAILV AAPQFMETKO KENECIGDYP EVLOEIWVPL RVNETNLEGF
 201 LLPLIMSYC YFRIOITLES CNHKKAKAI KLILVIVF FLEWFPYNYM
 251 IFLETLKIYD FFPSCDMRKD LRLALSVET VAFSHCLNP LIYAFAGEKF
 301 RRYIYHYGK CLAVIGRSV HYDFSSSSQ RSRHGSVLS NRYHNSDD
 351 ALLLL

11AA SEQUENCE 1.0

ID W54728 standard; peptide: 24 AA.
 AC W54728:
 DT 25-SEP-1998 (first entry)
 DE Peptide from 11 p80-103.
 KW Mannose; antigen; antigen-presenting cell; mannosylated peptide; T cell;
 KW vaccine; treatment.
 OS Synthetic.
 PN MO9813378-A1.
 PD 02-APR-1998.
 PE 25-SEP-1997: NL0536.
 PR 26-SEP-1996: EP-202701.
 PA (UYLE-) RIJSGUNIV LEIDEN.
 PI Drijfhout JW, Koning F;
 DR WPI: 98-230631/20.
 PT Increasing uptake and presentation of antigen(s) - by adding mannose
 PT residue(s) to antigen for increasing T cell response, useful in,
 PT e.g. vaccines against viral infection(s)
 PS Disclosure: Page 34, 47pp; English.
 CC The peptides W54539-W54809 are examples of peptides to which at least 1
 CC (preferably 2) mannose can be attached to increase their uptake as
 CC antigens by antigen-presenting cells. Uptake of agonist mannosylated
 CC peptides will increase the T cell response, whereas uptake of antagonist
 CC peptides blocks the T cell response. Blocking binding of immunogenic
 CC autoantigens can be used in treatment of type I diabetes, rheumatoid
 CC arthritis, graft rejection etc., also to induce T-cell non-
 CC responsiveness. Vaccines containing mannosylated antigen are used to
 CC prevent or treat infections by, e.g. bacteria, viruses, fungi, helminths
 CC and parasites.
 SQ Sequence 24 AA;
 W54728 Length: 24 February 11, 2000 15:48 Type: P Check: 3126 ..

1 LKPPKPYSK MRWATPLMQ ALPM
 ID W60258 standard; Protein: 339 AA.
 AC W60258:
 DT 28-SEP-1998 (first entry)
 DE Klebsiella pneumoniae glycerol-3-phosphate dehydrogenase.
 KW glycerol-3-phosphate dehydrogenase; production;
 KW Klebsiella pneumoniae.
 OS WO9821341-A2.
 PN 22-MAY-1998.
 PR 13-NOV-1997: U20873.
 PR 13-NOV-1996: US-030601.
 PA (GENV) GENENCOR INT INC.
 PI Chase MW, Diaz-Corres M, Dunn-Coleman NS, Timbur D;
 DR WPI: 98-297944/26.
 PT New method for increasing production of 1,3-propane:diol - comprises
 PT fermentation of inexpensive carbon sources by microorganism
 PT expressing dehydratase, used, e.g. to prolong half-life of enzyme
 PS Disclosure: Page 70-71; 13pp; English.
 CC The sequence is that of glycerol-3-phosphate dehydrogenase.
 CC It was used as part of a method of fermentative production
 CC of 1,3-propanediol (1,3-pd) using an organism comprising
 CC at least 1 gene encoding a dehydratase, is improved by
 CC inserting into the host a gene encoding protein X and culturing
 CC the transformant in presence of a carbon source (e.g. mono-, oligo-
 CC or poly-saccharide or IC substrate) convertible to 1,3-pd.
 CC 1,3-pd is a starting material for polyesters, polyurethanes and
 CC cyclic compounds. 1,3-pd can now be produced by a single
 CC recombinant organism from inexpensive carbon sources such
 CC as glucose (rather than costly glycerol or dihydroxyacetone),
 CC rapidly and without causing pollution.
 SQ Sequence 339 AA;
 W60258 Length: 339 February 11, 2000 15:48 Type: P Check: 4443 ..

1 MNQRNASMTV IGAGSYCTAL ATLIARNGHE VLVGHPDEH IATLERDRCN
 51 AAFLPDVPF DTLHESDLA TALASRNIL VVPSHVGE VLQOIKPLMR

Mon Feb 14 08:07:18 2000

ags.cat

Page 67

101 PDARLYWATK GLEAFETGRLL QDVAREALSD QIPLAVISGP TFARKEIAGL
151 PTAISLASTD CTEFADDLQOL LHCKSFERY SNPDFIGVOL GGAANKVIAI
201 GAGSODIGF GANARTALIT RGLAENSRLG MALLGADPATF MGNAGLGLV
251 LTCIDNOSRN RRFQMLGOG MDVQSOEKI GQVEGYRNT KEVRELAHRE
301 GVEPIITEEL XOVIVGKNA REAALTLLGR ARKDESSH
11AA_SEQUENCE 1.0
ID W60267 standard: Protein: 387 AA.
AC W60267; 387-SEP-1998 (first entry)
DE Klebsiella pneumoniae DHAT protein.
KM glycerol-3-phosphate dehydratase; production;
OS Klebsiella pneumoniae.
PN W09821341-A2.
PD 22-MAY-1998.
PF 13-NOV-1997: U20873.
PR 13-NOV-1996: US-030601.
PI (GENV) GENECOR INT INC.
DR Chase MW, Diaz-Lorres M, Dunn-Coleman NS, Trimbur D;
PI WPI: 98-297944/26.
PT New method for increasing production of 1,3-propanediol - comprises
PT fermentation of inexpensive carbon sources by microorganism
PT expressing dehydratase, used, e.g. to prolong half-life of enzyme
PS Disclolase; Page 96-97; 133p; English.
CC The sequence is that of DHAT protein used as part of a method of
CC fermentative production of 1,3-propanediol (1,3-pd), using an organism
CC comprising at least 1 gene encoding a dehydratase, is improved by
CC inserting into the host a gene encoding protein X and culturing
CC the transformant in presence of a carbon source (e.g. mono-, oligo-
CC or poly-saccharide or 1C substrate) convertible to 1,3-pd.
CC 1,3-pd is a starting material for polyesters, polyurethanes and
CC cyclic compounds. 1,3-pd can now be produced by a single
CC recombinant organism from inexpensive carbon sources such
CC as glucose (rather than costly glycerol or dihydroxyacetone),
CC rapidly and without causing pollution.
SQ Sequence 387 AA;
W60267 Length: 387 February 11, 2000 15:48 Type: P Check: 843 ..
1 MSTRMPDIYV PNVNFFGNA ISVYGERCOL LGKKALLVT DKGLRAIKDG
51 AUKTILHYLR EAGIEVALFD GVEPNKNDN VRDGLAVRR EOCDIIVTG
101 GGSFPHDCKG IGIAATHRGD LYQYAGITEL TNPPIYAV NTAGTASEV
151 TRHCVLNTE TKKFEVYISW RKLPSVIND PLMIGRPA LTAATGMDL
201 THAVEAVISK DANPYTDAA MOLAIRIARN LQAVALGNS LQAREMAYA
251 SLAGNAFNN ANLGYVHAMA HQJGLYDMP HGVAANVLLP HVARYNLIN
301 PEKADIATL MGNITGLST LDAEKAIAA TIRLSMDIGI PQHRLDLYK
351 EADPYVAM ALKDNAFSN PRKNGOEIA AIFROAF
11AA_SEQUENCE 1.0
ID W60764 standard: Protein: 260 AA.
AC W60764;
DE 28-SEP-1998 (first entry)
DE Rainbow trout interleukin 1 beta.
KM Interleukin 1 beta; IL1; beta; vaccine; fish; medicine.
OS Oncorhynchus mykiss.
PN W09817802-A1.
PD 30-APR-1998.
PF 16-OCT-1997: G02855.
PR 06-NOV-1996: GB-023173.
PR 17-OCT-1996: GB-021681.
PR (UYAB) UNIV ABERDEEN.

PI Cunningham C, Secombes CJ, Zou J;
DR WPI: 98-261498/23.
DR N-PSDB: V36179.
PT Rainbow trout interleukin-1beta - can be administered as adjuvant in
PT existing and developing fish vaccines
PS Claim 14: Fig 3: 46pp; English.
CC The sequence is that of rainbow trout interleukin 1 (IL-1)
CC beta which can be used in the manufacture of a medication for fish
CC e.g. it can be used as a adjuvant in existing and developing vaccines.
SQ Sequence 260 AA;
W60764 Length: 260 February 11, 2000 15:48 Type: P Check: 7422 ..
1 MDESNYSLI KNTSESAMS SKLPQGLDE VSHHPITWRH IANLIAMER
51 LKGEQVYMG TEFKDKDLN FLESAVEEH IYLEESAPP ASRRAGFSS
101 TSQRECVTD SENKCVMLN ENAEIHAMML QGSSYHKVH LNLSTVTPV
151 PIETEARPA LDKSNNLT SCSKSGRPT LHLEVADKD QLKSIQOSD
201 MYRELFYRN TGVDSITLES ASFRNWFIST DMQDYTPV DMQKXAPNR
251 LTFEIORHN
11AA_SEQUENCE 1.0
ID W57889 standard: Protein: 271 AA.
AC W57889;
DE 23-SEP-1998 (first entry)
DE Corn raffinose synthetase.
KM Raffinose synthetase; metabolism modification; food additive;
KM gastrointestinal flora; corn.
OS Zea mays.
PN EP-849359-A2.
PD 24-JUN-1998.
PF 18-DEC-1997: 122417.
PR 18-DEC-1996: JP-338673.
PR (SUMO) SUMIOMO CHEM CO LTD.
PI Oeda K, Mantabe E;
DR WPI: 98-324670/29.
DR N-PSDB: V40803.
PT New nucleic acid molecule encoding plant raffinose synthetase -
PT capable of producing raffinose, used as food additives with
PT beneficial effects on gastrointestinal flora
PS Claim 1: Page 39-40; 44pp; English.
CC This sequence represents the corn raffinose synthetase of the
CC invention. The raffinose synthetase is capable of producing raffinose by
CC combining a D-galactosyl group through an alpha (1-6) bond with a
CC hydroxyl group attached to the carbon atom at position 6 of a D-glucose
CC residue in a sucrose molecule. The DNA can be used to modify metabolism
CC of a host organism by introducing into the host organism or its cell so
CC that the content of the raffinose family oligosaccharides in the host
CC organism or cell is changed. Raffinose oligosaccharides are useful as
CC food additives with beneficial effects on the gastrointestinal flora.
SQ Sequence 271 AA;
W57889 Length: 271 February 11, 2000 15:48 Type: P Check: 5290 ..
1 OSTRPCAFH AASRAISGP IYVSDVGH DFLAIRLAL PGCIYRCG
51 HALPRLCLF ADPLHDGRV LKIMVNRPA GVVAFCNCG GGNSEPARN
101 KCFSEFVPL AARASPDVE WRSGKAGPV SVKDVSOFAV YAVEARTLOL
151 LRPEGYDIL LQFTYELFV VAPRVISH RAIKAPIL ANNLNTAGAV
201 QAFARKDAS GYLAEVFKG AGEIVAYSSA TPRLCKVNGD EAEFTYKDV
251 VYDVFWSSG SSKLCQGYV Y
11AA_SEQUENCE 1.0
ID W44905 standard: Protein: 39 AA.
AC W44905;

DT 29-SEP-1998 (first entry)
DE "polyproline beta-turn helix" polypeptide Delta-PRO4-beta.
KM Chimeric; polypoline; beta-turn helix; facilitation; inhibition;
KM retrovirus; envelope glycoprotein; gene therapy; cystic fibrosis;
KM myopathy; lysosomal disorder; cancer; human immunodeficiency virus.
OS Synthetic.
PN WO9744474-A2.
PD 27-NOV-1997.
PF 16-MAY-1997; F00870.
PR 20-MAY-1996; FR-006234.
PI (CNRS) CENT NAT RECH SCI.
PI Cosset FL, Russell SJ, Valseesia S;
DR MPI: 98-018526/02.
DR N-PSDB: V30424.
PT Introducing genes into eukaryotic cells using peptide with two
PT receptor binding regions - especially as part of a viral
PT glyco-protein, used in gene therapy to ensure precise targeting of
PT vectors to particular cell types.
PS Disclosure: Fig 36; 117pp; French.
CC A novel chimeric protein for gene transfer comprises, at the N-terminus,
CC a protein binding a target molecule on the cell surface, and, at the
CC C-terminus, a second protein recognising a secondary target molecule on
CC the cell surface. The two domains are separated by a peptide comprising
CC at least 30% pro residues arranged so as to induce approximately 180 deg.
CC folds (beta or reverse turns) in the peptide sequence, the folds being
CC regularly spread to form a polyproline beta-turn helix structure. This
CC sequence is an example of a pro-rich linking peptide. The pro-rich
CC peptide acts by facilitating or inhibiting binding between the C-terminal
CC protein and its receptor, and inhibition occurs as long as the N-terminal
CC protein does not interact with its target. The whole construct especially
CC forms part of a (retro)viral envelope glycoprotein. The recombinant
CC (retro)viral particles are used for in vitro or ex vivo transfer of
CC nucleic acid into target cells, especially for gene therapy, e.g. in
CC cystic fibrosis, myopathy, lysosomal disorders, cancer, viral (especially
CC human immunodeficiency virus) infections, etc.
SQ Sequence 39 AA;
W44905 Length: 39 February 11, 2000 15:48 Type: P Check: 1135 ..
1 AAADPTIMFS LTRQVNLNCP RVPICPNVL PDORLPSSA
ID W44906 standard; Protein: 70 AA.
AC W44906;
DT 29-SEP-1998 (first entry)
DE "polyproline beta-turn helix" polypeptide Delta-PRO4-int.
KM Chimeric; polypoline; beta-turn helix; facilitation; inhibition;
KM retrovirus; envelope glycoprotein; gene therapy; cystic fibrosis;
KM myopathy; lysosomal disorder; cancer; human immunodeficiency virus.
OS Synthetic.
PN WO9744474-A2.
PD 27-NOV-1997.
PF 16-MAY-1997; F00870.
PR 20-MAY-1996; FR-006234.
PI (CNRS) CENT NAT RECH SCI.
PI Cosset FL, Russell SJ, Valseesia S;
DR MPI: 98-018526/02.
DR N-PSDB: V30425.
PT Introducing genes into eukaryotic cells using peptide with two
PT receptor binding regions - especially as part of a viral
PT glyco-protein, used in gene therapy to ensure precise targeting of
PT vectors to particular cell types.
PS Disclosure: Fig 37; 117pp; French.
CC A novel chimeric protein for gene transfer comprises, at the N-terminus,
CC a protein binding a target molecule on the cell surface, and, at the
CC C-terminus, a second protein recognising a secondary target molecule on
CC the cell surface. The two domains are separated by a peptide comprising
CC at least 30% pro residues arranged so as to induce approximately 180 deg.
CC folds (beta or reverse turns) in the peptide sequence, the folds being
CC regularly spread to form a polyproline beta-turn helix structure. This
CC sequence is an example of a pro-rich linking peptide. The pro-rich
CC peptide acts by facilitating or inhibiting binding between the C-terminal
CC protein and its receptor, and inhibition occurs as long as the N-terminal
CC protein does not interact with its target. The whole construct especially
CC forms part of a (retro)viral envelope glycoprotein. The recombinant
CC (retro)viral particles are used for in vitro or ex vivo transfer of
CC nucleic acid into target cells, especially for gene therapy, e.g. in
CC cystic fibrosis, myopathy, lysosomal disorders, cancer, viral (especially
CC human immunodeficiency virus) infections, etc.
SQ Sequence 127 AA;
W44907 Length: 127 February 11, 2000 15:48 Type: P Check: 3740 ..
1 AAATGTQAV WKPISMDLI SLKRGNTPMW TGCSVACGP CYDSKXNS
ID W44908 standard; Protein: 79 AA.
AC W44908;
DT 29-SEP-1998 (first entry)
DE "polyproline beta-turn helix" polypeptide PRO-beta.
KM Chimeric; polypoline; beta-turn helix; facilitation; inhibition;
KM retrovirus; envelope glycoprotein; gene therapy; cystic fibrosis;
KM myopathy; lysosomal disorder; cancer; human immunodeficiency virus.
OS Synthetic.
PN WO9744474-A2.
PD 27-NOV-1997.

CC protein does not interact with its target. The whole construct especially
CC forms part of a (retro)viral envelope glycoprotein. The recombinant
CC (retro)viral particles are used for in vitro or ex vivo transfer of
CC nucleic acid into target cells, especially for gene therapy, e.g. in
CC cystic fibrosis, myopathy, lysosomal disorders, cancer, viral (especially
CC human immunodeficiency virus) infections, etc.
SQ Sequence 70 AA;
W44906 Length: 70 February 11, 2000 15:48 Type: P Check: 6449 ..
1 AAANPLVIEF TDGKRWMD GPKSWGRLRY RTGDTPTIME SLTROVLNNG
ID W44907 standard; Protein: 127 AA.
AC W44907;
DT 29-SEP-1998 (first entry)
DE "polyproline beta-turn helix" polypeptide Delta-PRO4-vrb.
KM Chimeric; polypoline; beta-turn helix; facilitation; inhibition;
KM retrovirus; envelope glycoprotein; gene therapy; cystic fibrosis;
KM myopathy; lysosomal disorder; cancer; human immunodeficiency virus.
OS Synthetic.
PN WO9744474-A2.
PD 27-NOV-1997.
PF 16-MAY-1997; F00870.
PR 20-MAY-1996; FR-006234.
PI (CNRS) CENT NAT RECH SCI.
PI Cosset FL, Russell SJ, Valseesia S;
DR MPI: 98-018526/02.
DR N-PSDB: V30426.
PT Introducing genes into eukaryotic cells using peptide with two
PT receptor binding regions - especially as part of a viral
PT glyco-protein, used in gene therapy to ensure precise targeting of
PT vectors to particular cell types
PS Disclosure: Fig 38; 117pp; French.
CC A novel chimeric protein for gene transfer comprises, at the N-terminus,
CC a protein binding a target molecule on the cell surface, and, at the
CC C-terminus, a second protein recognising a secondary target molecule on
CC the cell surface. The two domains are separated by a peptide comprising
CC at least 30% pro residues arranged so as to induce approximately 180 deg.
CC folds (beta or reverse turns) in the peptide sequence, the folds being
CC regularly spread to form a polyproline beta-turn helix structure. This
CC sequence is an example of a pro-rich linking peptide. The pro-rich
CC peptide acts by facilitating or inhibiting binding between the C-terminal
CC protein and its receptor, and inhibition occurs as long as the N-terminal
CC protein does not interact with its target. The whole construct especially
CC forms part of a (retro)viral envelope glycoprotein. The recombinant
CC (retro)viral particles are used for in vitro or ex vivo transfer of
CC nucleic acid into target cells, especially for gene therapy, e.g. in
CC cystic fibrosis, myopathy, lysosomal disorders, cancer, viral (especially
CC human immunodeficiency virus) infections, etc.
SQ Sequence 127 AA;
W44907 Length: 127 February 11, 2000 15:48 Type: P Check: 3740 ..
1 AAATGTQAV WKPISMDLI SLKRGNTPMW TGCSVACGP CYDSKXNS
ID W44908 standard; Protein: 79 AA.
AC W44908;
DT 29-SEP-1998 (first entry)
DE "polyproline beta-turn helix" polypeptide PRO-beta.
KM Chimeric; polypoline; beta-turn helix; facilitation; inhibition;
KM retrovirus; envelope glycoprotein; gene therapy; cystic fibrosis;
KM myopathy; lysosomal disorder; cancer; human immunodeficiency virus.
OS Synthetic.
PN WO9744474-A2.
PD 27-NOV-1997.

PF 16-MAY-1997; F00870.
 PR 20-MAY-1996; FR-006234.
 PA (CNRS) CENT NAT RECH SCI.
 PI Cosset FL, Russell SJ, Valsesia S;
 DR WPI: 98-018526/02.
 DR N-PSDB: V30427.
 PT Introducing genes into eukaryotic cells using peptide with two
 PT receptor binding regions - especially as part of a viral
 PT glyco-protein, used in gene therapy to ensure precise targeting of
 PT vectors to particular cell types.
 PS Disclosure: Fig 39; 117pp; French.
 CC A novel chimeric protein for gene transfer comprises, at the N-terminus,
 CC a protein binding a target molecule on the cell surface, and, at the
 CC C-terminus, a second protein recognising a secondary target molecule on
 CC the cell surface. The two domains are separated by a peptide comprising
 CC at least 30% pro residues arranged so as to induce approximately 180 deg.
 CC folds (beta or reverse turns) in the peptide sequence, the folds being
 CC regularly spread to form a polypyrroline beta-turn helix structure. This
 CC sequence is an example of a pro-rich linking peptide. The pro-rich
 CC peptide acts by facilitating or inhibiting binding between the C-terminal
 CC protein and its receptor, and inhibition occurs as long as the N-terminal
 CC protein does not interact with its target. The whole construct especially
 CC forms part of a (retroviral) envelope glycoprotein. The recombinant
 CC (retroviral) particles are used for in vitro or ex vivo transfer of
 CC nucleic acid into target cells, especially for gene therapy, e.g. in
 CC cystic fibrosis, myopathy, lysosomal disorders, cancer, viral (especially
 CC human immunodeficiency virus) infections, etc.
 SQ Sequence 79 AA;

W44908 Length: 79 February 11, 2000 15:48 Type: P Check: 1205 ..

1 AAADPTMFS LFRVYLVNCP RVPICGNPVL PDQRLPSSPI EIVAPQPPS
 51 PLNTSTPST TSPSTSPS PSVPQPPA

!!AA-SEQUENCE 1.0
 ID W44909; standard; Protein: 110 AA.
 AC W44909;
 DE "29-SEP-1998 (first entry)
 DE "polypyrroline beta-turn helix" polypeptide PRO-int.
 DE "Chimeric; polypyrroline; beta-turn helix; facilitation; inhibition;
 DE "retrovirus; envelope glycoprotein; gene therapy; cystic fibrosis;
 DE "myopathy; lysosomal disorder; cancer; human immunodeficiency virus.
 OS Synthetic.
 PN MO9744474-A2.
 PD 27-NOV-1997.
 PR 16-MAY-1997; F00870.
 PR 20-MAY-1996; FR-006234.
 PA (CNRS) CENT NAT RECH SCI.
 PI Cosset FL, Russell SJ, Valsesia S;
 DR WPI: 98-018526/02.
 DR N-PSDB: V30428.
 PT Introducing genes into eukaryotic cells using peptide with two
 PT receptor binding regions - especially as part of a viral
 PT glyco-protein, used in gene therapy to ensure precise targeting of
 PT vectors to particular cell types.
 PS Disclosure: Fig 40; 117pp; French.
 CC A novel chimeric protein for gene transfer comprises, at the N-terminus,
 CC a protein binding a target molecule on the cell surface, and, at the
 CC C-terminus, a second protein recognising a secondary target molecule on
 CC the cell surface. The two domains are separated by a peptide comprising
 CC at least 30% pro residues arranged so as to induce approximately 180 deg.
 CC folds (beta or reverse turns) in the peptide sequence, the folds being
 CC regularly spread to form a polypyrroline beta-turn helix structure. This
 CC sequence is an example of a pro-rich linking peptide. The pro-rich
 CC peptide acts by facilitating or inhibiting binding between the C-terminal
 CC protein and its receptor, and inhibition occurs as long as the N-terminal
 CC protein does not interact with its target. The whole construct especially
 CC forms part of a (retroviral) envelope glycoprotein. The recombinant
 CC (retroviral) particles are used for in vitro or ex vivo transfer of
 CC nucleic acid into target cells, especially for gene therapy, e.g. in
 CC cystic fibrosis, myopathy, lysosomal disorders, cancer, viral (especially
 CC human immunodeficiency virus) infections, etc.

SQ Sequence 110 AA;

W44909 Length: 110 February 11, 2000 15:48 Type: P Check: 4699 ..

1 AAANPLYLEF TDAGKRNAD GPKSMGLRLY RGTGDTIME SLTQVNLVNG
 51 PRVPIGNPV LPDQRLPSSP IEIVAPQPP SPLNTSTPSS TTSPTSTSPST
 101 SPSTPQPPA

!!AA-SEQUENCE 1.0
 ID W44910; standard; Protein: 167 AA.
 AC W44910;
 DE 29-SEP-1998 (first entry)
 DE "polypyrroline beta-turn helix" polypeptide PRO-vrb.
 DE "Chimeric; polypyrroline; beta-turn helix; facilitation; inhibition;
 DE "retrovirus; envelope glycoprotein; gene therapy; cystic fibrosis;
 DE "myopathy; lysosomal disorder; cancer; human immunodeficiency virus.
 OS Synthetic.
 PN MO9744474-A2.
 PD 27-NOV-1997.
 PR 16-MAY-1997; F00870.
 PR 20-MAY-1996; FR-006234.
 PA (CNRS) CENT NAT RECH SCI.
 PI Cosset FL, Russell SJ, Valsesia S;
 DR WPI: 98-018526/02.
 DR N-PSDB: V30429.
 PT Introducing genes into eukaryotic cells using peptide with two
 PT receptor binding regions - especially as part of a viral
 PT glyco-protein, used in gene therapy to ensure precise targeting of
 PT vectors to particular cell types.
 PS Disclosure: Fig 41; 117pp; French.
 CC A novel chimeric protein for gene transfer comprises, at the N-terminus,
 CC a protein binding a target molecule on the cell surface, and, at the
 CC C-terminus, a second protein recognising a secondary target molecule on
 CC the cell surface. The two domains are separated by a peptide comprising
 CC at least 30% pro residues arranged so as to induce approximately 180 deg.
 CC folds (beta or reverse turns) in the peptide sequence, the folds being
 CC regularly spread to form a polypyrroline beta-turn helix structure. This
 CC sequence is an example of a pro-rich linking peptide. The pro-rich
 CC peptide acts by facilitating or inhibiting binding between the C-terminal
 CC protein and its receptor, and inhibition occurs as long as the N-terminal
 CC protein does not interact with its target. The whole construct especially
 CC forms part of a (retroviral) envelope glycoprotein. The recombinant
 CC (retroviral) particles are used for in vitro or ex vivo transfer of
 CC nucleic acid into target cells, especially for gene therapy, e.g. in
 CC cystic fibrosis, myopathy, lysosomal disorders, cancer, viral (especially
 CC human immunodeficiency virus) infections, etc.
 SQ Sequence 167 AA;

W44910 Length: 167 February 11, 2000 15:48 Type: P Check: 1990 ..

1 AAATGGQAV WKPTSSMDLI SLKRGNTPMD TGCEKVNACG CYLDSKVSNS
 51 FQAGTRGRC NPLVLEFQDA GKRANWDGPK SMGLRLYTG TDPITMFSLT
 101 RQVNLVNGPRV PIGPNVPLD QRLPSSPIEL VPAQPPSPL NTSPTSTTS
 151 TSPSTPSS VPQPPA

!!AA-SEQUENCE 1.0
 ID W44879; standard; Protein: 104 AA.
 AC W44879;
 DE 29-SEP-1998 (first entry)
 DE "polypyrroline beta-turn helix" polypeptide PRO-(4070A).
 DE "Chimeric; polypyrroline; beta-turn helix; facilitation; inhibition;
 DE "retrovirus; envelope glycoprotein; gene therapy; cystic fibrosis;
 DE "myopathy; lysosomal disorder; cancer; human immunodeficiency virus.
 OS Synthetic.
 PN MO9744474-A2.
 PD 27-NOV-1997.
 PR 16-MAY-1997; F00870.
 PR 20-MAY-1996; FR-006234.

PA (CNRS) CENT NAT RECH SCI.
 PI Cosset FL, Russell SJ, Valsesia S;
 DR WPI: 98-018526/02.
 N-PSDB: V30398.
 PT Introducing genes into eukaryotic cells using peptide with two
 PT receptor binding regions - especially as part of a viral
 PT glyco-protein, used in gene therapy to ensure precise targeting of
 PT vectors to particular cell types
 PS Disclosure: Fig 10: 117pp; French.
 CC A novel chimeric protein for gene transfer comprises, at the N-terminus,
 CC a protein binding a target molecule on the cell surface, and, at the
 CC C-terminus, a second protein recognising a secondary target molecule on
 CC the cell surface. The two domains are separated by a peptide comprising
 CC at least 304 pro. residues arranged so as to induce approximately 180 deg.
 CC folds (beta or reverse turns) in the peptide sequence, the folds being
 CC regularly spread to form a polypyrroline beta-turn helix structure. This
 CC sequence is an example of a Pro-rich linking peptide. The pro-rich
 CC peptide acts by facilitating or inhibiting binding between the C-terminal
 CC protein and its receptor, and inhibition occurs as long as the N-terminal
 CC forms part of a (retro)viral envelope glycoprotein. The recombinant
 CC (retro)viral particles are used for in vitro or ex vivo transfer of
 CC nucleic acid into target cells, especially for gene therapy, e.g. in
 CC cystic fibrosis, myopathy, lysosomal disorders, cancer, viral (especially
 CC human immunodeficiency virus) infections, etc.
 SQ Sequence 104 AA;
 W44879 Length: 104 February 11, 2000 15:48 Type: P Check: 9655 ..

1 EFTDAGKRN WDGPKSWGLR LYRTGDPIT MSLTRQVLN VGPRVPGPN
 51 PVLDPQRLPS SPLIIVAPQ PPSPLNTSTP PSTTSTPETS PTSPVPGPP
 101 PAA
 11AA-SEQUENCE 1.0
 ID W44882 standard; Protein: 61 AA.
 AC W44882;
 DT 29-SEP-1998 (first entry)
 DE "polypyrroline beta-turn helix" polypeptide Delta-PRO.
 KW Chimeric; polypyrroline; beta-turn helix; facilitation; inhibition;
 KW retrovirus; envelope glycoprotein; gene therapy; cystic fibrosis;
 KW myopathy; lysosomal disorder; cancer; human immunodeficiency virus.
 OS Synthetic.
 PN WO9744474-A2.
 PD 27-NOV-1997.
 PR 16-MAY-1997: F00870.
 PR 20-MAY-1996: FR-006234.
 PA (CNRS) CENT NAT RECH SCI.
 PI Cosset FL, Russell SJ, Valsesia S;
 DR WPI: 98-018526/02.
 N-PSDB: V30401.
 PT Introducing genes into eukaryotic cells using peptide with two
 PT receptor binding regions - especially as part of a viral
 PT glyco-protein, used in gene therapy to ensure precise targeting of
 PT vectors to particular cell types
 PS Disclosure: Fig 13: 117pp; French.
 CC A novel chimeric protein for gene transfer comprises, at the N-terminus,
 CC a protein binding a target molecule on the cell surface, and, at the
 CC C-terminus, a second protein recognising a secondary target molecule on
 CC the cell surface. The two domains are separated by a peptide comprising
 CC at least 304 pro. residues arranged so as to induce approximately 180 deg.
 CC folds (beta or reverse turns) in the peptide sequence, the folds being
 CC regularly spread to form a polypyrroline beta-turn helix structure. This
 CC sequence is an example of a Pro-rich linking peptide. The pro-rich
 CC peptide acts by facilitating or inhibiting binding between the C-terminal
 CC protein and its receptor, and inhibition occurs as long as the N-terminal
 CC forms part of a (retro)viral envelope glycoprotein. The recombinant
 CC (retro)viral particles are used for in vitro or ex vivo transfer of
 CC nucleic acid into target cells, especially for gene therapy, e.g. in
 CC cystic fibrosis, myopathy, lysosomal disorders, cancer, viral (especially
 CC human immunodeficiency virus) infections, etc.

SQ Sequence 61 AA;
 W44882 Length: 61 February 11, 2000 15:48 Type: P Check: 8918 ..
 1 EFTDAGKRN WDGPKSWGLR LYRTGDPIT MSLTRQVLN VGPRVPGPN
 51 PVLDPQRLPS
 11AA-SEQUENCE 1.0
 ID W62612 standard; Protein: 369 AA.
 AC W62612;
 DT 21-SEP-1998 (first entry)
 DE Human glutamate-binding protein (HGLUBP).
 KW Human; glutamate-binding protein; HGLUBP; Incyte clone 386116;
 KW treatment; prevention; diagnosis; central nervous system disorder;
 KW signalling; glutamate; ischaemia; hypoglycaemia; epilepsy; convulsion;
 KW schizophrenia; neurodegeneration; amyotrophic lateral sclerosis;
 KW lathyrism.
 OS Homo sapiens.
 FH Key.
 FH Region.
 FT Location/Qualifiers
 FT 178..21
 FT /note="transmembrane region"
 FT 211..230
 FT /note="transmembrane region"
 FT 260..285
 FT /note="transmembrane region"
 FT 297..320
 FT /note="transmembrane region"
 PN WO9821241-A1.
 PD 22-MAY-1998.
 PR 10-NOV-1997: U20860.
 PR 14-NOV-1996: US-749289.
 PA (INCITE) INCITE PHARM INC.
 PI Bandman O, Coleman R;
 DR WPI: 98-297888/26.
 N-PSDB: V38502.
 DT New human glutamate-binding protein - useful for treatment,
 DT prevention and diagnosis of central nervous system diseases, e.g.
 DT Alzheimer's
 PS Claim 1: Fig 1A-E; 66pp; English.
 CC The present sequence represents human glutamate-binding protein (HGLUBP).
 CC The protein has structural homology to the rat glutamate-binding subunit
 CC of the N-methyl-D-aspartate receptor. The nucleic acid encoding HGLUBP
 CC was first identified in Incyte clone 386116 from a thymus cDNA library.
 CC HGLUBP is useful for treatment, prevention and diagnosis of central
 CC nervous system disorders that require an increase or reduction in
 CC signalling through glutamate, e.g. ischaemia, hypoglycaemia, epilepsy,
 CC convulsions, dementia related to AIDS, schizophrenia, neurodegeneration
 CC such as Alzheimer's, Huntington's, Creutzfeldt-Jacob or Parkinson's
 CC diseases, amyotrophic lateral sclerosis or lathyrism, pain, defective
 CC motor rhythm, vasomotor tone, baroreception and long-term potentiation,
 CC muscle relaxation and sedation. Antibodies can also be used to target
 CC pharmaceuticals to HGLUBP-expressing cells or to diagnose diseases
 CC associated with HGLUBP expression and to monitor treatment.
 SQ Sequence 369 AA;
 W62612 Length: 369 February 11, 2000 15:49 Type: P Check: 8912 ..

1 MKAVFMCIGT TLIPPTDIR GGSHPDPPM LSLPLIGLIT HSPSSPPT
 51 VSGGTMAPA PTPKATHRV PTEKATHRG PYPQGYPOG PYPQGYPOG
 101 PYPQSPFPN PYGQDPVPG QDPDSDPHGN YQEEGPPSY DNOFPATNW
 151 DDKSIRQAVI RKVELVLTLQ LSVLTSTVSV FTFVAEKG FVREWWTYYV
 201 SAVFPSILI VLSCGDFFR KHPMNLVALS VLTASLSYV GMTASFYNT
 251 AVTMAVGIT AVCTVAIFS MÖTRYDFISC MGVLVSMV LFFAILCIP
 301 IRRNILEIYV ASLGALEFIC FLAVDTQLL GNKQSLSP EYFALNLV
 351 TIIINIFLYI LTIIGRAKE

11AA_SEQUENCE 1.0
 ID W30679 standard; Protein: 387 AA.
 AC M30679.
 DT 12-OCT-1998 (first entry)
 DE 1,3-propanediol oxidoreductase.
 KW 1,3-propanediol oxidoreductase; DHAP; 1,3-propanediol.
 OS Klebsiella pneumoniae ATCC 25955.
 PN M09821339-A1.
 PD 22-MAY-1998.
 PR 10-NOV-1997; U20292.
 PA (DUP) DU POINT DE NEUOURS & CO E I.
 PI GENEX) GENECOR INT INC.
 PI Dias-Torres M, Gatenby AA, Haynie SL, Hsu AK, Lareau RD,
 PI Nagarajan V, Nair RV, Nakamura CE, Payne MS, Picataggio SK,
 PI Tribur DE, Whited GM.
 DR WPI: 98-297942/26.
 DR N-PSDB: V42015 and V42022.
 PT Fermentative production of 1,3-propanediol - by single organism
 PT containing cassette comprising specific genes, and capable of using
 PT inexpensive carbon sources
 PS Claim 13; Page 81-82; 95pp; English.
 CC This is the 1,3-propanediol oxidoreductase encoded by the dhap
 CC gene (see V42015) of Klebsiella pneumoniae ATCC 25955. This
 CC enzyme catalyses the reduction of 3-hydroxypropionaldehyde to
 CC 1,3-propanediol (I). A claimed method for production of (I)
 CC comprises culturing a microorganism transformed with a cassette
 CC containing at least 1 of the genes (see V42012-21) for
 CC glycerol-3-phosphate dehydrogenase, glycerol-3-phosphatase,
 CC glycerol dehydratase and 1,3-propanediol oxidoreductase (see
 CC M30676-85). Any of these genes not present on the cassette must
 CC be present endogenously. The enzyme sequences may include
 CC substitutions, deletions and additions provided activity is not
 CC altered. A single recombinant organism can now be used for
 CC production of (I) from inexpensive C-sources (contrast use of
 CC glycerol or dihydroxyacetone) without causing pollution. (I) is
 CC a starting material for polyesters, polyurethanes and cyclic
 CC compounds. 387 AA:
 SQ Sequence 387 AA:
 W30679 Length: 387 February 11, 2000 15:49 Type: P Check: 843 ..

1 MSYRMEDYLV PNVNFECPNA ISYVGECOL LGKKALLIVT DKGLRAIKDG
 51 AVDKTLHYLR EAGIEVAIFD GVEPNKDTN VRDGLAVERR EOCDIIVTGG
 101 GSPHDCGKG IGIAATHEGD LYQAGIELT INPLPIYAV NTTAGTASEV
 151 TRHCYLINTE TKRFEIVISW RKLPSYSIND PLUMIGKPPA LTAATGMAL
 201 TRAVEYISK DANPVTDAAL MQAIRLIARN LROAVALGNS IQARENNAYA
 251 SLLAGAFNN ANLYGVHMA HQGGGLYDMP HGVANAVLLP HVARYMLIAN
 301 PEKFAIDIEL MGENITGLST LDAEKAIAA ITRLSMDIGI PQHRLDGVK
 351 EADFPYMAEM ALKQGNAFSN PRKGNQOEIA AIFRQAF

11AA_SEQUENCE 1.0
 ID W69232 standard; Protein: 263 AA.
 AC W69232.
 DT 20-OCT-1998 (first entry)
 DE Fc receptor-like protein; phagocytosis inducer; rheumatoid arthritis;
 KW immune complex related disease; systemic lupus erythematosus; allergy;
 KW haemolytic anaemia; thrombocytopenia; anaphylaxis; cancer; lymphoma;
 KW leukemia; infection; immunomodulator; viral entry inhibitor; therapy;
 KW FcR-II.
 OS Homo sapiens.
 FH key
 FT key
 FT Peptide
 FT Location/Qualifiers
 FT 1..18
 FT /note= "signal peptide"

FT Protein 19..263
 FT /note= "mature FcR-II"
 PN W09831806-A2.
 PD 23-JUL-1998.
 PR 20-JAN-1998; U01184.
 PR 18-JUN-1997; US-049872.
 PR 21-JAN-1997; US-034205.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Genz RL, Murphy M, Ni J, Olsen HS, Ruben SM;
 DR WPI: 98-414105/35.
 DR N-PSDB: V44825.
 PT Nucleic acid encoding Fc receptor-like polypeptides or their
 PT fragments and related vectors, transformed cells and antibodies,
 PT useful for treating and diagnosing diseases of the haematopoietic
 PT and immune systems
 PS Claim 23; Fig 2A, 14pp; English.
 CC This sequence is the Fc receptor-like II protein (FcR-II) of the
 CC invention. Cells containing the DNA are used to express the recombinant
 CC protein, and to screen for specific (ant)agonists. The proteins are used
 CC to induce phagocytosis, and their (ant)agonists are used to treat immune
 CC complex related diseases (e.g. rheumatoid arthritis, systemic lupus
 CC erythematosus, haemolytic anaemia, thrombocytopenia, anaphylaxis,
 CC allergy, colorectal or breast cancer, lymphoma, leukaemia, infection by
 CC intracellular pathogens etc). The antagonists are also useful as
 CC immunomodulators and inhibitors of viral (e.g. human immune deficiency or
 CC dengue viruses) entry into cells. The proteins may also be used to screen
 CC for specific binding agents, i.e. (ant)agonists, for raising antibodies
 CC (Ab), and for identification of particular cells or tissues. The Ab can
 CC be used therapeutically as antagonists; as assay reagents for diagnostic
 CC determination of the levels of expression of the proteins and for
 CC useful as hybridisation probes or primers for isolating related genes, in
 CC situ hybridisation (chromosome mapping) and diagnostically to measure
 CC mRNA expression.
 SQ Sequence 263 AA:
 W69232 Length: 263 February 11, 2000 15:49 Type: P Check: 3029 ..

1 MALVLIQLL TLMPCHTDI TSPVPASVH KPWLAGQPA TVVTPGVNVT
 51 LRCRAPQPM RGLFKPGEI APILFRDVS ELAEFFLEEV TPAGGSGYRC
 101 CYRRPMGPG WMSQPSDYLE LVTLELPR SLVALPGVV GPANSLRC
 151 AGRIRNMSV LYREGVAPL QYRHSQPMV DETLLGAPR GTYSCYHTP
 201 SAPTVSQRS EYLVTSWEDS GSSDYTRGNL VRLGIAGVL ISGALVTFD
 251 WRSQNRAPAG IRP

11AA_SEQUENCE 1.0
 ID W62617 standard; Protein: 212 AA.
 AC W62617.
 DT 27-OCT-1998 (first entry)
 DE Rattus norvegicus SOCS1 protein.
 KW SOCS: suppressor of cytokine signalling; PCR primer;
 KW autoimmunity disease; diagnosis; cancer; treatment;
 KW cytokine mediated cellular responsiveness; hyperimmunity;
 KW immunosuppression; allergies; hypertension.
 OS Rattus norvegicus.
 PN M09820023-A1.
 PD 14-MAY-1998.
 PR 31-OCT-1997; AU0729.
 PR 14-FEB-1997; AU-005117.
 PR 01-NOV-1996; AU-003384.
 PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
 PI Alexander WS, Hilton DJ, Metcalf D, Nicholson SE,
 PI Nicolson R, Richardson RT, Starr R, Viney EM, Willson TA;
 DR WPI: 98-286854/25.
 DR N-PSDB: V38863.
 PT Suppressor of cytokine signalling proteins - useful to treat
 PT disease, injury or abnormality involving cytokine mediated cellular
 PT responsiveness e.g. hyperimmunity, immunosuppression, allergies and

PT hypertension
PS Claim 13, Page 119-120: 3256p: English.
CC The sequence is that of a suppressor of cytokine signalling
CC protein (SOCS). SOCS can be used to screen for naturally
CC occurring antibodies to SOCS, which may occur, e.g. in some autoimmune
CC diseases. Alternatively, specific antibodies can be used to
CC screen for SOCS, which is useful as a knowledge of SOCS levels
CC may be important for the diagnosis of certain cancers. Soluble
CC SOCS polypeptides can be used to treat disease, injury or
CC abnormally involving cytokine mediated cellular responsiveness,
CC e.g. hyperinflammation, immunosuppression, allergies and hypertension.
SQ Sequence 212 AA;

CC protein (SOCS). SOCS can be used to screen for naturally
 CC occurring antibodies to SOCS, which may occur, e.g. in some autoimmune
 CC diseases. Alternatively, specific antibodies are used to
 CC screen for SOCS, which is useful as a knowledge of SOCS levels
 CC may be important for the diagnosis of certain cancers. Soluble
 CC SOCS polypeptides can be used to treat disease, injury or
 CC abnormality involving cytokine mediated cellular responsiveness,
 CC e.g. hyperimmunity, immunosuppression, allergies and hypertension.
 SQ Sequence 134 AA;

W62624 Length: 134 February 11, 2000 15:49 Type: P Check: 8778 ..

1 GGMWLGNNRL YHDKNOPSQ TYPFLEPDE TEIVDSFEV ALDMXQJLTS
 51 FIVDQYMGV AFRGLKGGKL YPVSAVWGH CEIRRYLNG LDPEPLMD
 101 LCRNSVRLAL GKRLGAIIPA LPLPASLKA Y LIXO

11AA-SEQUENCE 1.0
 ID W62613 standard; Protein: 212 AA.
 AC W62613;
 DT 27-OCT-1998 (first entry)
 DE Mus musculus SOCS1 protein.
 KW SOCS: suppressor of cytokine signalling; PCR primer;
 KW autoimmune disease; diagnosis; cancer; treatment;
 KW cytokine mediated cellular responsiveness; hyperimmunity;
 KW immunosuppression; allergies; hypertension.
 OS Mus musculus.
 PN WO9820023-A1.
 PD 14-MAY-1998.
 PE 31-OCT-1997; AU-00729.
 PR 14-FEB-1997; AU-005117.
 PR 01-NOV-1996; AU-003384.
 PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
 PI Alexander WS, Hilton DJ, Metcalf D, Nicholson SE,
 PI Nicola NA, Richardson RT, Starr R, Viney EM, Willson TA;
 PI WPI: 98-286854/25.
 DR N-PSDB: V38659.
 DR Suppressor of cytokine signalling proteins - useful to treat
 PT disease, injury or abnormality involving cytokine mediated cellular
 PT responsiveness e.g. hyperimmunity, immunosuppression, allergies and
 PT hypertension.
 PS Claim 13; Page 110; 325pp; English.
 CC The sequence is that of a suppressor of cytokine signalling
 CC protein (SOCS). SOCS can be used to screen for naturally
 CC occurring antibodies to SOCS, which may occur, e.g. in some autoimmune
 CC diseases. Alternatively, specific antibodies can be used to
 CC screen for SOCS, which is useful as a knowledge of SOCS levels
 CC may be important for the diagnosis of certain cancers. Soluble
 CC SOCS polypeptides can be used to treat disease, injury or
 CC abnormality involving cytokine mediated cellular responsiveness,
 CC e.g. hyperimmunity, immunosuppression, allergies and hypertension.
 SQ Sequence 212 AA;

W62613 Length: 212 February 11, 2000 15:49 Type: P Check: 4527 ..

1 MVARNOVAD NATSPAEP RRSPESSSS SSSPAAPVP RCPDAPAPA
 51 PGDHFRRFR SHSDYRRIR TSALIDACG YMGPLSVGA HERLRAPVG
 101 TEIVDSRQR NCFEALSVK ASGPTIRVH FQGRPHLDG SRETFDCLFE
 151 LLEHYAAR RMGAPLQR RVRPQELCR QRIYAVGRE NLRPIPLNV
 201 LRDYLSFPF QI

11AA-SEQUENCE 1.0
 ID W62614 standard; Protein: 198 AA.
 AC W62614;
 DT 27-OCT-1998 (first entry)
 DE Mus musculus SOCS2 protein.
 KW SOCS: suppressor of cytokine signalling; PCR primer;
 KW autoimmune disease; diagnosis; cancer; treatment;
 KW cytokine mediated cellular responsiveness; hyperimmunity;
 KW immunosuppression; allergies; hypertension.

W62614 Length: 198 February 11, 2000 15:49 Type: P Check: 1060 ..

1 MTRCLPSPG NGADRTSRQW GTAGLPEPQS PPARLAKAL RELSQTGW
 51 GSNVNEAKE KLEAPEGTF LINDSHSDY LITIVKTSQ GPINLRLEVO
 101 DKFPLDSII CVKSKIKQFD SYVALIDYV QMCKDKRTGP EAPRNGTVHL
 151 YTRKFLYISA PTIQHFCRLA INKQTGIVG LPLPTRLADY LEEYKFOV

11AA-SEQUENCE 1.0
 ID W62615 standard; Protein: 225 AA.
 AC W62615;
 DT 27-OCT-1998 (first entry)
 DE Mus musculus SOCS3 protein.
 KW SOCS: suppressor of cytokine signalling; PCR primer;
 KW autoimmune disease; diagnosis; cancer; treatment;
 KW cytokine mediated cellular responsiveness; hyperimmunity;
 KW immunosuppression; allergies; hypertension.
 OS Mus musculus.
 PN WO9820023-A1.
 PD 14-MAY-1998.
 PE 31-OCT-1997; AU-005117.
 PR 14-FEB-1997; AU-003384.
 PR 01-NOV-1996; AU-003384.
 PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
 PI Alexander WS, Hilton DJ, Metcalf D, Nicholson SE,
 PI Nicola NA, Richardson RT, Starr R, Viney EM, Willson TA;
 PI WPI: 98-286854/25.
 DR N-PSDB: V38661.
 DR Suppressor of cytokine signalling proteins - useful to treat
 PT disease, injury or abnormality involving cytokine mediated cellular
 PT responsiveness e.g. hyperimmunity, immunosuppression, allergies and
 PT hypertension.
 PS Claim 13; Page 114-115; 325pp; English.
 CC The sequence is that of a suppressor of cytokine signalling
 CC protein (SOCS). SOCS can be used to screen for naturally
 CC occurring antibodies to SOCS, which may occur, e.g. in some autoimmune
 CC diseases. Alternatively, specific antibodies can be used to
 CC screen for SOCS, which is useful as a knowledge of SOCS levels
 CC may be important for the diagnosis of certain cancers. Soluble
 CC SOCS polypeptides can be used to treat disease, injury or
 CC abnormality involving cytokine mediated cellular responsiveness,
 CC e.g. hyperimmunity, immunosuppression, allergies and hypertension.

W62615 Length: 225 February 11, 2000 15:49 Type: P Check: 1060 ..

1 MTRCLPSPG NGADRTSRQW GTAGLPEPQS PPARLAKAL RELSQTGW
 51 GSNVNEAKE KLEAPEGTF LINDSHSDY LITIVKTSQ GPINLRLEVO
 101 DKFPLDSII CVKSKIKQFD SYVALIDYV QMCKDKRTGP EAPRNGTVHL
 151 YTRKFLYISA PTIQHFCRLA INKQTGIVG LPLPTRLADY LEEYKFOV

11AA-SEQUENCE 1.0
 ID W62616 standard; Protein: 225 AA.
 AC W62616;
 DT 27-OCT-1998 (first entry)
 DE Mus musculus SOCS3 protein.
 KW SOCS: suppressor of cytokine signalling; PCR primer;
 KW autoimmune disease; diagnosis; cancer; treatment;
 KW cytokine mediated cellular responsiveness; hyperimmunity;
 KW immunosuppression; allergies; hypertension.
 OS Mus musculus.
 PN WO9820023-A1.
 PD 14-MAY-1998.
 PE 31-OCT-1997; AU-005117.
 PR 14-FEB-1997; AU-003384.
 PR 01-NOV-1996; AU-003384.
 PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
 PI Alexander WS, Hilton DJ, Metcalf D, Nicholson SE,
 PI Nicola NA, Richardson RT, Starr R, Viney EM, Willson TA;
 PI WPI: 98-286854/25.
 DR N-PSDB: V38661.
 DR Suppressor of cytokine signalling proteins - useful to treat
 PT disease, injury or abnormality involving cytokine mediated cellular
 PT responsiveness e.g. hyperimmunity, immunosuppression, allergies and
 PT hypertension.
 PS Claim 13; Page 114-115; 325pp; English.
 CC The sequence is that of a suppressor of cytokine signalling
 CC protein (SOCS). SOCS can be used to screen for naturally
 CC occurring antibodies to SOCS, which may occur, e.g. in some autoimmune
 CC diseases. Alternatively, specific antibodies can be used to
 CC screen for SOCS, which is useful as a knowledge of SOCS levels
 CC may be important for the diagnosis of certain cancers. Soluble
 CC SOCS polypeptides can be used to treat disease, injury or
 CC abnormality involving cytokine mediated cellular responsiveness,
 CC e.g. hyperimmunity, immunosuppression, allergies and hypertension.

```

SQ Sequence 225 AA:
W62615 Length: 225 February 11, 2000 15:49 Type: P Check: 9828
1 MYH5KFPAA GMSRPLDTSL RKTFSKSE YOLVYNAVAK LQSGFYWSA
51 YVGGENALL SAEPAGTFLI RDSSDGRHFF TLVSATQSGT KNIRIQCEGG
101 SFSLSQDPSS TQVPREDCV LKLVHYMPP PGTPEFSLPP TEESSEVPEQ
151 PPAQALPGST PKRAYIYSG GKKIPLYLR PLSSNVATLQ HLCKRTVNGH
201 LQSEKVTQL PGPIREFLDQ YDAPL
!!AA_SEQUENCE 1.0
ID W62616 standard; Protein: 211 AA.
AC W62616;
DT 27-OCT-1998 (first entry)
DE Homo sapiens SOCS1 protein.
KW SOCS; suppressor of cytokine signaling; PCR primer;
KW autoimmune disease; diagnosis; cancer; treatment;
KW cytokine mediated cellular responsiveness; hyperimmunity;
KW immunosuppression; allergies; hypertension.
OS Homo sapiens.
PN WO9820023-A1.
PD 14-MAY-1998.
PF 31-OCT-1997; AU0729.
PR 14-FEB-1997; AU-005117.
PR 01-NOV-1996; AU-003384.
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
PI Alexander WS, Hilton DJ, Metcalf D, Nicholson SE.
PI Nicola NA, Richardson RT, Starr R, Viney EW, Willson TA.
PI WPI: 98-268654/25.
DR N-PSDB: V38682.
PT Suppressor of cytokine signaling proteins - useful to treat
PT disease, injury or abnormality involving cytokine mediated cellular
PT responsiveness e.g. hyperimmunity, immunosuppression, allergies and
PT hypertension
PS Claim 13: Page 116-117; 325pp; English.
CC The sequence is that of a suppressor of cytokine signaling
CC protein (SOCS). SOCS can be used to screen for naturally
CC occurring antibodies to SOCS, which may occur. e.g. in some autoimmune
CC diseases. Alternatively, specific antibodies can be used to
CC screen for SOCS, which is useful as a knowledge of SOCS levels
CC may be important for the diagnosis of certain cancers. Soluble
CC SOCS polypeptides can be used to treat disease, injury or
CC abnormality involving cytokine mediated cellular responsiveness,
CC e.g. hyperimmunity, immunosuppression, allergies and hypertension.
SQ Sequence 211 AA:
W62616 Length: 211 February 11, 2000 15:49 Type: P Check: 758
1 MVAHNQVAD NAVSTAEPK RPPESSSS SSPAPAPR PCPAPAPAP
51 GDHFTFFS HADRRITRA SALLDACGY WGPLSVGHAH ERLRAPVGT
101 FLVDRDRGN CFPAISVYMA SGPTSTRVHF QAGRHLDS RESFDCLFEL
151 LEHYVAAPR MGAPLRORR VRPLQELCRQ RIVATGREN LARIPLNPVL
201 RDYLSFPQ I
!!AA_SEQUENCE 1.0
ID W61619 standard; Protein: 250 AA.
AC W61619;
DT 27-OCT-1998 (first entry)
DE Clone HRP86 of TM4SF superfamily.
KW Human; receptor; immune disorder; cancers; blood disorder;
KW juvenile rheumatoid arthritis; Graves disease.
OS Homo sapiens.
PN WO9831799-A2.
PD 23-JUL-1998.
PF 21-JAN-1998; U00959.
21-JAN-1997; US-034205.
PR 21-JAN-1997; US-034204.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Gentz RL, Ni J, Rosen CA.
PI WPI: 98-427559/36.
DR N-PSDB: V48114.
PT New isolated polynucleotide(s) and encoded receptor polypeptide(s)
PT - used to develop products for diagnosing or treating e.g. immune
PT disorders, cancers, blood disorders or immuno-compromised disease
PT states
PS Claim 11: Page 38-39; 79pp; English.
CC Clone HRP86 is a member of the TM4SF receptor superfamily. The
CC products generated using the receptor can be used for treating abnormal
CC conditions related to both an excess of and insufficient amounts of
CC receptor activity. They can be used in the treatment of e.g. immune
CC disorders, cancers, blood disorders, juvenile rheumatoid arthritis,
CC Graves disease or immunocompromised disease states. The products can
CC also be used for detection and diagnosis.
SQ Sequence 250 AA:
W61619 Length: 250 February 11, 2000 15:49 Type: P Check: 1162
1 MNSMTAVPV ANSVLVAPH NGIYVTPGIM SHVPLPNSQ PQVHLVGNP
51 PSLVSNVNGQ PVQKALREK TLGAIQIIG LAHIGISIM ATVLVGEYLS
101 ISFYGFPEFW GGLWFLISGS LSVAAENQPY SYCLLSGSLG LNIYSAICSA
151 VGVILFITDL SIPHYAVPD YYPYAMGVNP GMAISGVLLV FCLLEFGIAC
201 ASHFQGLV CQSSNVSVI YPNYIANPV IPEPYTSP SYSEIQAANK
!!AA_SEQUENCE 1.0
ID W65460 standard; Protein: 375 AA.
AC W65460;
DT 09-NOV-1998 (first entry)
DE Human growth differentiation factor-8.
KW Growth differentiation factor-8; GDF-8; human.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Modified_site 71
FT /note= "N-glycosylated"
FT 263..266
FT /note= "RxxR proteolytic cleavage site"
PN WO9835019-A1.
PD 13-AUG-1998.
PF 06-FEB-1997; US-795671.
PR 06-FEB-1997; US-795671.
PA (UYUO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
PI Lee S, McPherron AC.
PI WPI: 98-447217/38.
PT Transgenic animal growth differentiation factor-11 is inhibited - by
PT insertion of transgene, also use of GDF-11 inhibitors for treating
PT muscular wasting, neuromuscular disease, obesity
PS Example 3: Page 55-56; 89pp; English.
CC This is the amino acid sequence of human growth differentiation
CC factor-8 (GDF-8). It shows a high degree of sequence homology
CC to the newly identified human growth differentiation factor-11
CC (GDF-11, see W65458). Alignment of the GDF-8 and GDF-11 sequences
CC reveals potential N-linked glycosylation signals and putative
CC proteolytic processing sites at analogous positions. The 2
CC sequences are related not only in the C-terminal region following
CC the putative cleavage site (90% amino acid sequence identity) but
CC also in the pro-region of the molecules (45% amino acid sequence
CC identity. Claimed transgenic animals in which GDF-11 production is
CC reduced produce higher than normal levels of muscle and are useful
CC in the food industry. GDF-11 polypeptides, polynucleotides and
CC antibodies can be used to modulate GDF-11 activity or gene
CC expression for treatment of cell proliferative disorders involving
CC muscle, nerve and adipose tissue.
SQ Sequence 375 AA:
W65460 Length: 375 February 11, 2000 15:49 Type: P Check: 1814
```

Mon Feb 14 08:07:18 2000

ags.cat

Page 75

1 MOKLQCVYI YLFMLIVAGP VDLNSENQK ENVEKEGICN ACTWRONTKS
51 SRLEAKIKOI LSKRLLETAP NISKDVIRQL LPKAPPLREL IDQYDVQRDD
101 SSGSLSEDDO YHATTEITIT MPESDEPLMO VDGKPKCOFF KESSKIOYK
151 VKAQLMILY RPVEPTTVE VOLRLIKPM KDGRIYIGIR SLKLDNPGT
201 GIMQSIDVKT VLQNMKQPE SNIGITIKAL DENGHDLAVT FPBGSDGLN
251 PLEKAVKDT PKRSRDEGL DCDHESTER CCRPLIYDF EAGGMWITA
301 PKRYKANCY GECEVFLQK YPHTLVHQA NPGSAGPCC TPKKSPINM
351 LYFNKEQIIT YKIPAMVVD RCGCS
11AA-SEQUENCE 1.0
ID W64373 standard; Protein; 120 AA.
AC W64373;
DE 09-NOV-1998 (first entry)
KW Mycobacterium tuberculosis antigen RD1P5.
OS Mycobacterium tuberculosis strain Erdman.
PN W09816645-A2.
PD 23-APR-1998.
PF 07-OCT-1997; U18214.
PR 13-MAR-1997; US-81811.
PR 11-OCT-1996; US-729622.
PA (CORI-) CORIYA CORP.
PI Reed SG, Skelly YAM, Twardik DR, Vedvick TS;
DR N-PSMB; V4442.
DR WPI: 98-251292/22.
PT New isolated Mycobacterium tuberculosis polyepitopes and DNA - used
PT to develop products for the detection of M. tuberculosis infection
PS and diagnosis of tuberculosis
PS Example 3; Page 203204-250pp; English
CC This polypeptide comprises an antigenic portion of Mycobacterium
CC tuberculosis antigen RD1P5. Its amino acid sequence was deduced
CC from a DNA molecule (see V44442) isolated from a M. tuberculosis
CC strain Erdman genomic DNA expression library using rabbit antisera
CC raised against M. tuberculosis fractionated proteins. Recombinant
CC RD1P5 was shown to stimulate cell proliferation and interferon
CC gamma production in T cells from M. tuberculosis-immune donors.
CC The invention relates to compositions and methods for diagnosing
CC tuberculosis. It provides polypeptides (see W64291-W64379)
CC comprising antigenic or immunogenic portions of M. tuberculosis
CC antigens, as well as DNA sequences encoding such polypeptides,
CC recombinant expression vectors and transformed or transfected host
CC cells. Also claimed are methods and diagnostic kits for detecting
CC M. tuberculosis infection in a patient using these polypeptides,
CC antibodies or oligonucleotide probes and primers.
SQ Sequence 120 AA;
W64373 Length: 120 February 11, 2000 15:49 Type: P Check: 3575 ..
1 TPERSFVDDL DIDSLSWEI AVQTEDEKYG KIPDEDLAQL RTGVDVAVYI
51 OKLEENPEEA AQALRAKIES ENPDARADR CVSPTSQARD ARRLARSAR
101 LACRRLPASV PTTTRDPRE
11AA-SEQUENCE 1.0
ID W38580 standard; Protein; 174 AA.
AC W38580;
DE 06-NOV-1998 (first entry)
KW Streptococcus pneumoniae protein of unknown function.
KW Streptococcus pneumoniae protein; genetic immunisation; antagonist;
KW Immunological response; inoculation; antibody production; inhibitor;
KW T cell immune response; antimicrobial compound; bacterial adhesion;
KW extracellular matrix protein; protein-mediated cell invasion; wound;
KW pathogenesis.

OS Streptococcus pneumoniae.
FH Key Location/Qualifiers
FT MISC_difference 1.174
FT /note- "residues designated X are unspecified, and
FT represented as Xaa in the specification"
PN W09743303-A1.
PD 20-NOV-1997.
PF 14-MAY-1997; U07950.
PR 14-MAY-1996; US-017670.
PA (SMIR) SMITHKLINE BEECHAM CORP.
PA (SMIR) SMITHKLINE BEECHAM PLC.
PI Black MT, Hodgson JE, Knowles DJC, Nicholas RO,
PI Shodola RK;
DR N-PSDB; T98632.
DR WPI: 98-008793/01.
PT Novel Streptococcus pneumoniae proteins and related DNA - useful for
PT diagnosing anti-microbial agents for treatment of bacterial
PT infections
PS Claim 12; Pages 351-352; 483pp; English.
CC This sequence represents a Streptococcus pneumoniae protein of
CC unknown function, and is encoded by a DNA sequence of the invention.
CC The DNA sequences, were isolated from Streptococcus pneumoniae strain
CC 0100993 (NCIMB 40794). The Streptococcus pneumoniae proteins of the
CC invention can be used to identify compounds which interact with and
CC inhibit or activate the activity of the proteins. Antagonists can be
CC used to treat diseases caused by S. pneumoniae proteins, through genetic
CC immunisation. They can also be used to induce an immunological response
CC in a mammal by inoculation with the S. pneumoniae proteins or delivery
CC of the encoding nucleic acids in a vector adequate to produce antibody
CC and/or T cell immune responses to protect the animal from disease. The
CC proteins can also be used to identify antimicrobial compounds which are
CC capable of inhibiting their bioactivity. In particular the proteins of
CC the invention can be used to prevent adhesion of bacteria to mammalian
CC extracellular matrix proteins on in-dwelling devices or in wounds, to
CC block protein-mediated mammalian cell invasion, and to block the normal
CC progression of pathogenesis in infections initiated other than by the
CC implantation of in-dwelling devices or other surgical techniques.
SQ Sequence 174 AA;
W38580 Length: 174 February 11, 2000 15:49 Type: P Check: 5404 ..
1 MGFNQPLSML ILHAKGLTES FIKSAQGTGA DRTTYISCNV RNHGAVDILK
51 YQELGYELAK VQPVDFXQT HHVETVALLS KLDVVDHISV EIEDEMDLT
101 SAEKATYIAQ IKRYVANKFE LKYSTLYIAQ IKKKOIEIR EHYKXKKDK
151 QIIPQCTPEK EEALMDALX FKMI
11AA-SEQUENCE 1.0
ID W70962 standard; Protein; 212 AA.
AC W70962;
DE 30-OCT-1998 (first entry)
DE A STAT function regulatory protein designated SIIS-1.
KW SIIS-1; STAT-induced inhibitor; STAT function;
KW JAK/STAT signal transmission system; STAT3; STAT6; Inhibit;
KW tyrosine phosphorylation; gp130; cytokine-regulatory protein; CIS;
KW screen; cytokine regulatory; inhibitory activity.
OS Mus sp.
PN W09830688-A1.
PD 16-JUL-1998.
PF 23-OCT-1997; J03860.
PR 10-JAN-1997; JP-014737.
PA (KISHU) KISHIMOTO T.
PI Naka T;
DR WPI: 98-399137/34.
DR N-PSDB; V42701.
PT STAT function regulatory protein - used in screening candidate
PT substances for cytokine regulatory activity
PS Claim 2; Pages 39-41; 60pp; Japanese.
CC The present sequence represents a protein (called SIIS-1, STAT-induced
CC inhibitor of STAT function 1) which regulates STAT protein function in
CC the JAK/STAT signal transmission system in mammalian cells. The protein

CC 1s induced by STAT3 or STAT6. It inhibits tyrosine phosphorylation of
CC STAT3 and of gp130. The STAT-1 protein sequence contains an SH2 domain
CC and is related to the cytokine-regulating protein CIS. S1S-1, or
CC transfectant cells expressing it, may be used to screen candidate
CC substances for cytokine regulatory or inhibitory activity.
SQ Sequence 212 AA;

W70962 Length: 212 February 11, 2000 15:49 Type: P Check: 4556 ..

1 MYARNOVAD NAISPAEPR RREPESSSS SSSPAAPVP RPQGVAPA
51 PGDTFRFR SHSDYRITR TSALDAGF YMGPLSVGA HEHLRAEPVG
101 TELVDSKOP NCFPLSVKM ASGPTSRVH FQAGFHLDG NRETFCLDE
151 LLEHYAAR RMGAPLROR RVRPDELOR QRYAAYGRE NLARILNPY
201 LRDYLSSEPF QI
!!AA_SEQUENCE 1.0
ID W77666 standard; Protein; 100 AA.
AC W77666;
DE 30-OCT-1998 (first entry)
KW Staphylococcus aureus protein of unknown function.
KW Staphylococcus aureus protein; immune response induction; eye infection;
KW antibody production; T-cell immune response; gastrointestinal infection;
KW respiratory infection; inhibitor; bacterial infection; cardiac infection;
KW central nervous system; kidney infection; urinary tract infection;
KW antimicrobial compound identification; broad spectrum antibiotic;
KW Staphylococcus aureus.
OS Key Location/Qualifiers
FH Misc_difference 4
FT Misc_difference 6 /note= "unspecified, encoded by CNT"
FT Misc_difference 8 /note= "unspecified, encoded by ANA"
FT Misc_difference 12 /note= "unspecified, encoded by GCN"
FT Misc_difference 42 /note= "unspecified, encoded by GCN"
FT Misc_difference 45 /note= "unspecified, encoded by GCN"
FT Misc_difference 45 /note= "unspecified, encoded by GNA"
EP-841394-A2.
PD 13-MAY-1998.
PR 24-SEP-1997; 307485.
PR 24-SEP-1996; US-027032.
PA (SMIR) SMITHKLINE BEECHAM CORP.
PI (SMIR) SMITHKLINE BEECHAM PLC.
PI Black MT, Burnham MKR, Hodgson JE, Knowles DUC,
PI Lonetto MA, Nicholas RO, Pratt JM, Reichard RW, Rosenberg M,
PI Ward JM; 98-252940/23.
DR N-PSDB: V53459
DR New nucleic acid sequences from Staphylococcus aureus WCHU29 -
PR useful in vaccines and for treatment of bacterial infections of e.g.
PR respiratory tract and central nervous system
PS Claim 11; Page 317-318; 390pp; English.
CC This sequence represents a DNA sequence of the invention.
CC function, and is encoded by a DNA sequence of the invention.
CC (NCIMB 40771). Host cells containing the DNA sequences are used to
CC produce polypeptides or fragments. The proteins are used in the treatment
CC of disease, for inducing an immune response by administering them, to
CC produce antibody and/or T-cell immune response. Antagonists of the
CC proteins are used for the inhibition of bacterial polypeptides.
CC Conditions which may be treated include bacterial infections, especially
CC respiratory, cardiac, gastrointestinal, central nervous, eye, kidney,
CC urinary tract, skin, bones and joints. The proteins can also be used to
CC identify antimicrobial compounds which are broad spectrum antibiotics,
CC especially useful in the treatment of H. pylori infection.
SQ Sequence 100 AA;

W77666 Length: 100 February 11, 2000 15:49 Type: P Check: 5825 ..

1 VMPYVEMKPF PXFIFTALP LAGVATGAI HFTANEYIPI GMLXNNGLI
51 AINLAYONLD RAFVODGTNI ESKLSLAATP KLASKGAIRE SIRLAIGANN
!!AA_SEQUENCE 1.0
ID W38716 standard; Protein; 161 AA.
AC W38716;
DE 10-NOV-1998 (first entry)
KW S. pneumoniae aspartyl tRNA synthetase.
KW Streptococcus pneumoniae protein; genetic immunisation; antagonist;
KW immunological response; inoculation; antibody production; inhibitor;
KW T cell immune response; antimicrobial compound; bacterial adhesion;
KW extracellular matrix protein; protein-mediated cell invasion; wound;
KW pathogenesis.
OS Streptococcus pneumoniae.
PW WP9743303-A1.
PD 20-NOV-1997.
PR 14-MAY-1997; U07950
PR 14-MAY-1996; US-017670.
PA (SMIR) SMITHKLINE BEECHAM CORP.
PI (SMIR) SMITHKLINE BEECHAM PLC.
PI Black MT, Hodgson JE, Knowles DUC, Nicholas RO,
PI Stodola RK; 98-008793/01.
DR N-PSDB: 198758.
DR Novel Streptococcus pneumoniae proteins and related DNA - useful for
PR diagnosing anti-microbial agents for treatment of bacterial
PR infections
PS Claim 12; Pages 450-451; 483pp; English.
CC This sequence represents a Streptococcus pneumoniae protein that, based
CC on homology with a Thermus aquaticus protein, is a aspartyl tRNA
CC synthetase, and is encoded by a DNA sequence of the invention.
CC The DNA sequences were isolated from Streptococcus pneumoniae strain
CC 0100993 (NCIMB 40794). The Streptococcus pneumoniae proteins of the
CC invention can be used to identify compounds which interact with and
CC inhibit or activate the activity of the proteins. Antagonists can be
CC used to treat diseases caused by S. pneumoniae proteins, through genetic
CC immunisation. They can also be used to induce an immunological response
CC in a mammal by inoculation with the S. pneumoniae proteins or delivery
CC of the encoding nucleic acids in a vector adequate to produce antibody
CC and/or T cell immune responses to protect the animal from disease. The
CC proteins can also be used to identify antimicrobial compounds which are
CC capable of inhibiting their bioactivity. In particular the proteins of
CC the invention can be used to prevent adhesion of bacteria to mammalian
CC extracellular matrix proteins on in-dwelling devices or in wounds, to
CC block protein-mediated mammalian cell invasion, and to block the normal
CC progression of pathogenesis in infections initiated other than by the
CC implantation of in-dwelling devices or other surgical techniques.
SQ Sequence 161 AA;

W38716 Length: 161 February 11, 2000 15:49 Type: P Check: 3773 ..

1 MKYDANALY GSKRPTRRD MLQDLTEV KYVDKFESE ALAAVAIVN
51 GLEPITANV SRKDIDKME VAKOGAGGL AMKAYVDGEL NCPVAKFLTG
101 IOEELITLTA LERKDYLYV ADLEEVANAT LGALRGRIAK EGLINDNKE
151 NFWLVDMWPM F
!!AA_SEQUENCE 1.0
ID W79302 standard; Protein; 67 AA.
AC W79302;
DE 24-NOV-1998 (first entry)
KW Staphylococcus aureus protein.
KW A. Staphylococcus aureus protein.
KW Alkaligenes eutrophus; cation efflux system czcd protein; treatment;
KW prevention; bacterial infection; Helicobacter pylori; vaccine.
OS Staphylococcus aureus.
PW WO9823738-A2.
PD 04-JUN-1998.
PR 24-NOV-1997; U22092.
PR 25-NOV-1996; US-031469.

PA (SMK) SMITHKLINE BEECHAM CORP.
 PT Warren RL;
 DR WPI: 98-322718/28.
 DR N-PSDB: V59870.
 PA New nucleic acid from *Staphylococcus aureus* NCIMB 40771 - useful
 PT for, e.g. diagnosis, prevention and treatment of bacterial
 PT infection(s).
 PT Claim 5, page 33: 114pp; English.
 CC The present sequence represents *Staphylococcus aureus* WCHU (NCIMB 40771)
 CC protein that has homology to the Alcaligenes eutrophus cation efflux
 CC system ccd protein. The protein is used to generate antibodies and to
 CC screen for antimicrobials. The products are used to treat or prevent
 CC bacterial infections, particularly where caused by *S. aureus* but also
 CC against *Helicobacter pylori*. Particular applications are to treat
 CC (subjects before surgery or insertion of an in-dwelling device
 CC (alternatively the device itself is impregnated before placement). The
 CC nucleic acid sequence is used as sources of antisense sequences (for
 CC therapeutic use) or regulatory elements for controlling expression of
 CC bacterial genes, and for antibacterial screening. The protein can be
 CC also used as a vaccine.
 CC Sequence 67 AA:
 SQ
 W79302 Length: 67 February 11, 2000 15:49 Type: P Check: 8790
 1 SHNINRGAF LHVIGDLGS VCAITRAILI WAFGWTADP IASILVSII
 51 LKSAWGITS SINIME
 IIAA_SEQUENCE 1.0
 ID W69889 standard; Protein: 376 AA.
 AC W69889;
 DT 07-DEC-1998 (first entry)
 DE Rat growth differentiation factor-8.
 KW Growth differentiation factor-8; GDF-8; rat; transgenic animal;
 KW transforming growth factor-beta; muscle; meat; inhibitor; obesity;
 KW neuromuscular disease; muscular dystrophy; cachexia; AIDS; cancer;
 KW therapy.
 OS *Rattus* sp.
 FH Key site Location/Qualifiers
 FT Cleavage_site 264..267
 FT Protein 268..376
 FT /label= Mat_protein
 FT W09833887-A1.
 PN 06-AUG-1998.
 PD 05-FEB-1998; 002479.
 PF 05-FEB-1998; US-862445.
 PR 05-FEB-1997; US-795071.
 PR 05-FEB-1997; US-795071.
 PR 28-APR-1997; US-847910.
 PA (UYJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
 PI Lee S, McPherson AC;
 DR WPI: 98-437444/37.
 DR N-PSDB: V45820.
 PT Transgenic animals with gene for growth differentiation factor-8
 PT disrupted - have increased muscle and reduced cholesterol contents,
 PT also use of GDF-8 inhibitors for treating cancer, obesity,
 PT neuromuscular disease
 PT Example 9; Fig 14d; 125pp; English.
 PS This is the amino acid sequence of rat growth differentiation
 CC factor-8 (GDF-8), a novel member of the transforming growth factor-
 CC beta superfamily that appears to relate to various cell
 CC proliferative disorders, especially those involving muscle, nerve
 CC and adipose tissue. The sequence was deduced from a cDNA clone
 CC (see V45820) isolated from a skeletal muscle cDNA library. The
 CC invention provides novel mammalian and avian GDF-8 proteins (see
 CC W69883-92). A transgenic non-human animal is claimed in which
 CC GDF-8 expression is disrupted or interfered with. Also claimed
 CC are: (1) chicken or turkey eggs or meat, beef, milk, pork and lamb
 CC animals by administering an antibody (Ab) that binds to GDF-8; (3)
 CC inhibiting the action of GDF-8 by treating foetal or adult muscle
 CC or progenitor cells with a GDF-8 inhibitor; (4) isolated nucleic
 CC acid encoding a GDF-8 protein truncated by loss of the C-terminal
 CC active fragment. The transgenic animals have increased muscle mass

CC and for poultry reduced cholesterol contents. Method (3) is used
 CC to treat muscle wasting or neuromuscular diseases, muscular atrophy
 CC and aging, particularly muscular dystrophy, spinal cord or
 CC traumatic injuries, congestive or obstructive lung disease, AIDS
 CC and cachexia. Method (4) is used to treat cancer of muscle,
 CC connective tissue and bone, or obesity. Also (not claimed) GDF-8
 CC can be used to maintain myoblasts intended for transplanting or to
 CC improve efficiency of fusion. Ab can be used to detect and
 CC quantify GDF-8 (particularly in muscle, for diagnosis or monitoring),
 CC also for immunotherapy and in vivo imaging.
 CC Sequence 376 AA:
 SQ
 W69889 Length: 376 February 11, 2000 15:49 Type: P Check: 1496
 1 MGRPMATV IYFVLVLAG PVLDNEDSR EANVEKEGIC NACAWRONTR
 51 YSRIFAIRIQ ILSKRLERA PNISKDAIRQ LPRAPPLRE LIQYDVQRD
 101 DSSDGLSD DDYATETETII TMPESDFLM QADGPKCF EFESSKIQYN
 151 KYRQAQMTV LRAVKPTTV FVQLRLIKP MMDGTRYTGI RSLKIDMSPG
 201 TGWOSIDVK TVLQNMKQP ESNIGIEIKA LDENGHDIAV TFPGGEDGL
 251 NFFLEKATD TPKRRRDFG LDODEHSTES RCRRYFLYVD FEARGWMI
 301 APRRYKANVC SGCEEFVFLQ KIPRHLYHQ ANRGSAGRC CTPTKMSPIN
 351 MLFNGKEQEI IYGRIPAMV DRGCS
 IIAA_SEQUENCE 1.0
 ID W69891 standard; Protein: 375 AA.
 AC W69891;
 DT 07-DEC-1998 (first entry)
 DE Pig growth differentiation factor-8.
 KW Growth differentiation factor-8; GDF-8; pig; transgenic animal;
 KW transforming growth factor-beta; muscle; meat; inhibitor; obesity;
 KW neuromuscular disease; muscular dystrophy; cachexia; AIDS; cancer;
 KW therapy.
 OS *Sus* scrofa.
 FH Key site Location/Qualifiers
 FT Cleavage_site 263..266
 FT Protein 267..375
 FT /label= Mat_protein
 FT W09833887-A1.
 PN 06-AUG-1998.
 PD 05-FEB-1998; 002479.
 PF 05-FEB-1998; US-862445.
 PR 05-FEB-1997; US-795071.
 PR 05-FEB-1997; US-795071.
 PR 28-APR-1997; US-847910.
 PA (UYJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
 PI Lee S, McPherson AC;
 DR WPI: 98-437444/37.
 DR N-PSDB: V45822.
 PT Transgenic animals with gene for growth differentiation factor-8
 PT disrupted - have increased muscle and reduced cholesterol contents,
 PT also use of GDF-8 inhibitors for treating cancer, obesity,
 PT neuromuscular disease
 PT Example 9; Fig 14f; 125pp; English.
 PS This is the amino acid sequence of porcine growth differentiation
 CC factor-8 (GDF-8), a novel member of the transforming growth factor-
 CC beta superfamily that appears to relate to various cell
 CC proliferative disorders, especially those involving muscle, nerve
 CC and adipose tissue. The sequence was deduced from a cDNA clone
 CC (see V45822) isolated from a skeletal muscle cDNA library. The
 CC invention provides novel mammalian and avian GDF-8 proteins (see
 CC W69883-92). A transgenic non-human animal is claimed in which
 CC GDF-8 expression is disrupted or interfered with. Also claimed
 CC are: (1) chicken or turkey eggs or meat, beef, milk, pork and lamb
 CC animals by administering an antibody (Ab) that binds to GDF-8; (3)
 CC inhibiting the action of GDF-8 by treating foetal or adult muscle
 CC or progenitor cells with a GDF-8 inhibitor; (4) isolated nucleic

CC acid encoding a GDF-8 protein truncated by loss of the C-terminal
 CC active fragment. The transgenic animals have increased muscle mass
 CC and for poultry reduced cholesterol contents. Method (3) is used
 CC to treat muscle wasting or neuromuscular diseases, muscular atrophy
 CC and aging, particularly muscular dystrophy, spinal cord or
 CC traumatic injuries, congestive or obstructive lung disease, AIDS
 CC and cachexia. Method (4) is used to treat cancer of muscle,
 CC connective tissue and bone, or obesity. Also (not claimed) GDF-8
 CC can be used to maintain myoblasts intended for transplanting or to
 CC improve efficiency of fusion. Ab can be used to detect and
 CC quantify GDF-8 (particularly in muscle, for diagnosis or monitoring),
 CC also for immunotherapy and in vivo imaging.

Sequence 375 AA;
 W69891 Length: 375 February 11, 2000 15:49 Type: P Check: 1805 ..

1 MOKIOTIYVI YLFMLIYAGP VDLNENSEOK ENVEKEGICN ACWNRONTSS
 51 SRLFAIKIQI LSKRLLETAP NISKDAIROL LPKAPPLREL IDQYDVQRDD
 101 SSDGLEDD YHATTEIIT MPTESDLMO VEGKPCCF KESSKIQYNK
 151 VKAQLMIYL RPVKPTTFV VOILRLIKPM KQTRTYGIR SLKLDNPGI
 201 GIMOSIDVKT VLOQMLKQPE SNLGEIKAL DENGHDIAVT FPEGEGEGLN
 251 PLEKAVTDT PKRSRDFGL DCDHESTESR CCRYPLTVDF EAFGMDIILA
 301 PKRYKAYCS GECEFLQK YPHTHLVQQA NPGSAGPCC TPTKMSPINM
 351 LYFNKEQIIT YGKIPAVVD RCGCS

11AA-SEQUENCE 1.0
 ID W69892 standard; Protein: 375 AA.

AC W69892.07-DEC-1998 (first entry)
 DT Ovine growth differentiation factor-8;
 DE Growth differentiation factor-8; GDF-8; sheep; transgenic animal;
 KW transforming growth factor-beta; muscle; meat; inhibitor; obesity;
 KM neuromuscular disease; muscular dystrophy; cachexia; AIDS; cancer;
 OS therapy.
 OS Ovis aries.
 FH Key
 FT Cleavage_site 263..266
 FT Protein 267..375
 FT /Label= Mat_protein
 PN W09833887-A1.
 PD 06-AUG-1998.
 PF 05-FEB-1998; U02479.
 PR 23-MAY-1997; US-862445.
 PR 05-FEB-1997; US-795071.
 PR 28-APR-1997; US-847910.
 PA (UJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
 PI Lee S, McPherson AC;
 DR WPI: 98-437444/37.
 DR N-PSDB: V45823.
 PT Transgenic animals with gene for growth differentiation factor-8
 PT also use of GDF-8 inhibitors for treating cancer, obesity,
 PT disrupted - have increased muscle and reduced cholesterol contents,
 PT also use of GDF-8 inhibitors for treating cancer, obesity,
 PT neuromuscular disease
 PS Example 9; Fig 14f; 125pp; English.
 CC This is the amino acid sequence of sheep growth differentiation
 CC factor-8 (GDF-8), a novel member of the transforming growth factor-
 CC beta superfamily that appears to relate to various cell
 CC proliferative disorders, especially those involving muscle, nerve
 CC and adipose tissue. The sequence was deduced from a cDNA clone
 CC (see V45823) isolated from a skeletal muscle cDNA library. The
 CC invention provides novel mammalian and avian GDF-8 proteins (see
 CC W69883-92). A transgenic non-human animal is claimed in which
 CC GDF-8 expression is disrupted or interfered with. Also claimed
 CC are: (1) chicken or turkey eggs or meat, beef, milk, pork and lamb
 CC from these animals; (2) method for increasing muscle mass in
 CC animals by administering an antibody (Ab) that binds to GDF-8; (3)

CC inhibiting the action of GDF-8 by treating foetal or adult muscle
 CC or progenitor cells with a GDF-8 inhibitor; (4) isolated nucleic
 CC acid encoding a GDF-8 protein truncated by loss of the C-terminal
 CC active fragment. The transgenic animals have increased muscle mass
 CC and for poultry reduced cholesterol contents. Method (3) is used
 CC to treat muscle wasting or neuromuscular diseases, muscular atrophy
 CC and aging, particularly muscular dystrophy, spinal cord or
 CC traumatic injuries, congestive or obstructive lung disease, AIDS
 CC and cachexia. Method (4) is used to treat cancer of muscle,
 CC connective tissue and bone, or obesity. Also (not claimed) GDF-8
 CC can be used to maintain myoblasts intended for transplanting or to
 CC improve efficiency of fusion. Ab can be used to detect and
 CC quantify GDF-8 (particularly in muscle, for diagnosis or monitoring),
 CC also for immunotherapy and in vivo imaging.

Sequence 375 AA;
 W69892 Length: 375 February 11, 2000 15:49 Type: P Check: 1229 ..

1 MOKIOTIYVI YLFMLIYAGP VDLNENSEOK ENVEKEGICN ACWNRONTSS
 51 SRLFAIKIQI LSKRLLETAP NISKDAIROL LPKAPPLREL IDQYDVQRDD
 101 SSDGLEDD YHATTEIIT MPTESDLMO VEGKPCCF KESSKIQYNK
 151 VKAQLMIYL RPVKPTTFV VOILRLIKPM KQTRTYGIR SLKLDNPGI
 201 GIMOSIDVKT VLOQMLKQPE SNLGEIKAL DENGHDIAVT FPEGEGEGLN
 251 PLEKAVTDT PKRSRDFGL DCDHESTESR CCRYPLTVDF EAFGMDIILA
 301 PKRYKAYCS GECEFLQK YPHTHLVQQA NPGSAGPCC TPTKMSPINM
 351 LYFNKEQIIT YGKIPAVVD RCGCS

11AA-SEQUENCE 1.0
 ID W69892 standard; Protein: 376 AA.

AC W69892.07-DEC-1998 (first entry)
 DT Murine growth differentiation factor-8;
 DE Growth differentiation factor-8; GDF-8; mouse; transgenic animal;
 KW transforming growth factor-beta; muscle; meat; inhibitor; obesity;
 KM neuromuscular disease; muscular dystrophy; cachexia; AIDS; cancer;
 OS therapy.
 OS Mus sp.
 FH Key
 FT Modified_site 72..74
 FT Cleavage_site 264..267
 FT Protein 268..376
 FT /Label= Mat_protein
 PN W09833887-A1.
 PD 06-AUG-1998.
 PF 05-FEB-1998; U02479.
 PR 23-MAY-1997; US-862445.
 PR 05-FEB-1997; US-795071.
 PR 28-APR-1997; US-847910.
 PA (UJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
 PI Lee S, McPherson AC;
 DR WPI: 98-437444/37.
 DR N-PSDB: V42113.
 PT Transgenic animals with gene for growth differentiation factor-8
 PT also use of GDF-8 inhibitors for treating cancer, obesity,
 PT disrupted - have increased muscle and reduced cholesterol contents,
 PT also use of GDF-8 inhibitors for treating cancer, obesity,
 PT neuromuscular disease
 PS Example 3; Fig 5a; 125pp; English.
 CC This is the amino acid sequence of mouse growth differentiation
 CC factor-8 (GDF-8), a novel member of the transforming growth factor-
 CC beta superfamily that appears to relate to various cell
 CC proliferative disorders, especially those involving muscle, nerve
 CC and adipose tissue. The sequence was deduced from a cDNA clone
 CC (see V42113) isolated from a skeletal muscle cDNA library. The
 CC invention provides novel mammalian and avian GDF-8 proteins (see
 CC W69883-92). A transgenic non-human animal is claimed in which

CC GDF-8 expression is disrupted or interfered with. Also claimed
 CC are: (1) chicken or turkey eggs or meat, beef, milk, pork and lamb
 CC from these animals; (2) method for increasing muscle mass in
 CC animals by administering an antibody (Ab) that binds to GDF-8; (3)
 CC inhibiting the action of GDF-8 by treating foetal or adult muscle
 CC or progenitor cells with a GDF-8 inhibitor; (4) isolated nucleic
 CC acid encoding a GDF-8 protein truncated by loss of the C-terminal
 CC active fragment. The transgenic animals have increased muscle mass
 CC and for poultry reduced cholesterol contents. Method (3) is used
 CC to treat muscle wasting or neuromuscular diseases, muscular atrophy
 CC and aging, particularly muscular dystrophy, spinal cord or
 CC traumatic injuries, congestive or obstructive lung disease, AIDS
 CC and cachexia. Method (4) is used to treat cancer of muscle,
 CC connective tissue and bone, or obesity. Also (not claimed) GDF-8
 CC can be used to maintain myoblasts intended for transplanting or to
 CC improve efficiency of fusion. Ab can be used to detect and
 CC quantify GDF-8 (particularly in muscle, for diagnosis or monitoring),
 CC also for immunotherapy and in vivo imaging.
 SQ Sequence 376 AA;

W30689 Length: 376 February 11, 2000 15:49 Type: P Check: 2293 ..

1 MMOKLQKIVY IYLFMLIAG PYDLNNGSR EENVEKGLC NMCARONTR

51 YRIEAIKIQ ILSKRLLETA PNISKDAIRO LPRAPPLRE LIDQDVQRD

101 DSDSDSLDD DYHATTEIT TMTESDFLM QADGKPCCF FFFSKIQYN

151 KYVKAQMIY LRPKPTPTV FVQILRLIKP MKDSTRYGI RSLKIDMSPG

201 TGIQSIDVK TVLQNLKOP ESNLGEIKA LDENGHDLAV TTPGGEDEL

251 NFLEVXTD TPKRSRDFG LDCDEHSTES RCRPYLTVD FFAFGMDWII

301 AKRKYKAVNC SGECEVFELQ KYPHHVLVHQ ANPRGSAGPC CPPTKMSPIN

351 MEYFKGKEGI YGKIPAMVY DRGCGS

11AA SEQUENCE 1.0

ID M69885 standard; Protein: 375 AA.

AC M69885;

DT 07-DEC-1998 (first entry)

DE Human growth differentiation factor-8.

KW Growth differentiation factor-8; GDF-8; human; transgenic animal;
 KW transforming growth factor-beta; muscle; meat; inhibitor; obesity;
 KW neuromuscular disease; muscular dystrophy; cachexia; AIDS; cancer;
 KW therapy.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Modified_site 71..73

FT FT /note="Asn is N-glycosylated"

FT FT Cleavage_site 263..266

FT FT Protein 267..375

FT FT /label=Mat_protein

PN MO9833887-A1.

PD 06-AUG-1998.

PF 05-FEB-1998; U02479.

PF 23-MAY-1997; US-862445.

PR 05-FEB-1997; US-795071.

PR 28-APR-1997; US-847910.

PA (UYJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.

PI Lee S, McPherson AC;

PI WPI: 98-437444/37.

DR N-PSDB: V45813.

PT Transgenic animals with gene for growth differentiation factor-8
 PT disrupted - have increased muscle and reduced cholesterol contents,
 PT also use of GDF-8 inhibitors for treating cancer, obesity,
 PT neuromuscular disease

PT Example 3; Fig 5c; 125pp; English.

PS This is the amino acid sequence of human growth differentiation
 CC factor-8 (GDF-8), a novel member of the transforming growth factor-
 CC beta superfamily that appears to relate to various cell
 CC proliferative disorders, especially those involving muscle, nerve

CC and adipose tissue. The sequence was deduced from a cDNA clone
 CC (see V45810) isolated from a skeletal muscle cDNA library. The
 CC invention provides novel mammalian and avian GDF-8 proteins (see
 CC M69883-92). A transgenic non-human animal is claimed in which
 CC GDF-8 expression is disrupted or interfered with. Also claimed
 CC are: (1) chicken or turkey eggs or meat, beef, milk, pork and lamb
 CC from these animals; (2) method for increasing muscle mass in
 CC animals by administering an antibody (Ab) that binds to GDF-8; (3)
 CC inhibiting the action of GDF-8 by treating foetal or adult muscle
 CC or progenitor cells with a GDF-8 inhibitor; (4) isolated nucleic
 CC acid encoding a GDF-8 protein truncated by loss of the C-terminal
 CC active fragment. The transgenic animals have increased muscle mass
 CC and for poultry reduced cholesterol contents. Method (3) is used
 CC to treat muscle wasting or neuromuscular diseases, muscular atrophy
 CC and aging, particularly muscular dystrophy, spinal cord or
 CC traumatic injuries, congestive or obstructive lung disease, AIDS
 CC and cachexia. Method (4) is used to treat cancer of muscle,
 CC connective tissue and bone, or obesity. Also (not claimed) GDF-8
 CC can be used to maintain myoblasts intended for transplanting or to
 CC improve efficiency of fusion. Ab can be used to detect and
 CC quantify GDF-8 (particularly in muscle, for diagnosis or monitoring),
 CC also for immunotherapy and in vivo imaging.
 SQ Sequence 375 AA;

M69885 Length: 375 February 11, 2000 15:49 Type: P Check: 1814 ..

1 MOKLQCYTI YLFMLVAGP VDLNENSEQK ENVEKEGLCN ACTWRONTRS

51 SRIEAIKIQ ILSKRLLETA PNISKDAIRO LPRAPPLREL IDQDVQRD

101 DSDSDSLDD YHATTEIT TMTESDFLMQ VDGKPCCF FFFSKIQYNK

151 VYKAQMIY LRPKPTPTV FVQILRLIKP MKDSTRYGI RSLKIDMSPG

201 TGIQSIDVK TVLQNLKOP ESNLGEIKA LDENGHDLAV TTPGGEDEL

251 NFLEVXTD TPKRSRDFG LDCDEHSTES RCRPYLTVD FFAFGMDWII

301 AKRKYKAVNC SGECEVFELQ KYPHHVLVHQ ANPRGSAGPC CPPTKMSPIN

351 MEYFKGKEGI YGKIPAMVY DRGCGS

11AA SEQUENCE 1.0

ID M69886 standard; Protein: 375 AA.

AC M69886;

DT 07-DEC-1998 (first entry)

DE Baboon growth differentiation factor-8.

KW Growth differentiation factor-8; GDF-8; baboon; transgenic animal;
 KW transforming growth factor-beta; muscle; meat; inhibitor; obesity;
 KW neuromuscular disease; muscular dystrophy; cachexia; AIDS; cancer;
 KW therapy.

OS Papio sp.

FT Key Location/Qualifiers

FT Modified_site 263..266

FT FT Cleavage_site 267..375

FT FT Protein /label=Mat_protein

PN MO9833887-A1.

PD 06-AUG-1998.

PF 05-FEB-1998; U02479.

PF 23-MAY-1997; US-862445.

PR 05-FEB-1997; US-795071.

PR 28-APR-1997; US-847910.

PA (UYJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.

PI Lee S, McPherson AC;

PI WPI: 98-437444/37.

DR N-PSDB: V45817.

PT Transgenic animals with gene for growth differentiation factor-8
 PT disrupted - have increased muscle and reduced cholesterol contents,
 PT also use of GDF-8 inhibitors for treating cancer, obesity,
 PT neuromuscular disease

PT Example 9; Fig 14a; 125pp; English.

PS This is the amino acid sequence of baboon growth differentiation
 CC factor-8 (GDF-8), a novel member of the transforming growth factor-

CC beta superfamily that appears to relate to various cell
 CC proliferative disorders, especially those involving muscle, nerve
 CC and adipose tissue. The sequence was deduced from a cDNA clone
 CC (see V45817) isolated from a skeletal muscle cDNA library. The
 CC invention provides novel mammalian and avian GDF-8 proteins (see
 CC W69883-92). A transgenic non-human animal is claimed in which
 CC GDF-8 expression is disrupted or interfered with. Also claimed
 CC are: (1) chicken or turkey eggs or meat, beef, milk, pork and lamb
 CC from these animals; (2) method for increasing muscle mass in
 CC animals by administering an antibody (Ab) that binds to GDF-8; (3)
 CC inhibiting the action of GDF-8 by treating foetal or adult muscle
 CC or progenitor cells with a GDF-8 inhibitor; (4) isolated nucleic
 CC acid encoding a GDF-8 protein truncated by loss of the C-terminal
 CC active fragment. The transgenic animals have increased muscle mass
 CC and for poultry reduced cholesterol contents. Method (3) is used
 CC to treat muscle wasting or neuromuscular diseases, muscular atrophy
 CC and aging, particularly muscular dystrophy, spinal cord or
 CC traumatic injuries, congestive or obstructive lung disease, AIDS
 CC and cachexia. Method (4) is used to treat cancer of muscle,
 CC connective tissue and bone, or obesity. Also (not claimed) GDF-8
 CC can be used to maintain myoblasts intended for transplanting or to
 CC improve efficiency of fusion. Ab can be used to detect and
 CC quantify GDF-8 (particularly in muscle, for diagnosis or monitoring),
 CC also for immunotherapy and in vivo imaging.
 CC Sequence 375 AA;

W69886 Length: 375 February 11, 2000 15:49 Type: P Check: 1463 ..

```

1 MOKIOLCVYI YLFMLIVAGP VDLNENSEQK ENVEKEGICN ACQWRONTKS
51 SRLFAIKIOI LSKRLLETAP NISKDAIROL LPKAPPLREL IDQYDVORD
101 SSDGSLDDDD YHATETIIT MPTESDLMO VDGKPKCCF KESSKIQYKN
151 VYKAOLMIYL RPYETPTTF VOILRLIKPM KDGRTYTGIR SLKIDNPGT
201 GIWOSIDVKT VLOWMLKQPE SNLGEIRKAL DENGHDLAVT FPPGEGDLN
251 PLEVKYVDT PKRSRDFGL DCDHSTESR CCRYPLTVDF EALGWDIITA
301 PKRYKANCYS GECEVFLQK YPHTILVHOA NPGSAGBPC TPTKMSPINM
351 LYFNKGEOII YGKIPAMVVD RCGCS

```

!!AA_SEQUENCE 1.0
 ID W69887 standard; Protein: 375 AA.
 AC W69887;
 DT 07-DEC-1998 (first entry)
 DE Bovine growth differentiation factor-8; transgenic animal;
 KW Growth differentiation factor-8; GDF-8; human; transgenic animal;
 KW transforming growth factor-beta; muscle; meat; inhibitor; obesity;
 KW neuromuscular disease; muscular dystrophy; cachexia; AIDS; cancer;
 OS therapy.
 FH key.
 FT cleavage_site 267..375
 FT Protein /label=Mat-Protein
 PN W09833887-A1.
 PD 06-AUG-1998;
 PE 05-FEB-1998; U02479.
 PR 23-MAY-1997; US-862445.
 PR 05-FEB-1997; US-795072.
 PR 28-APR-1997; US-847810.
 PA (UYUO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
 PI Lee S; McPherson AC;
 DR WPI: 98-437444/37.
 DR N-PSDB: V45818.
 PT Transgenic animals with gene for growth differentiation factor-8
 PT disrupted - have increased muscle and reduced cholesterol contents,
 PT also use of GDF-8 inhibitors for treating cancer, obesity,
 PT neuromuscular disease
 PS Example 9, Fig 14d, 125p: English.

CC This is the amino acid sequence of bovine growth differentiation
 CC factor-8 (GDF-8), a novel member of the transforming growth factor-
 CC beta superfamily that appears to relate to various cell
 CC proliferative disorders, especially those involving muscle, nerve
 CC and adipose tissue. The sequence was deduced from a cDNA clone
 CC (see V45818) isolated from a skeletal muscle cDNA library. The
 CC invention provides novel mammalian and avian GDF-8 proteins (see
 CC W69883-92). A transgenic non-human animal is claimed in which
 CC GDF-8 expression is disrupted or interfered with. Also claimed
 CC are: (1) chicken or turkey eggs or meat, beef, milk, pork and lamb
 CC from these animals; (2) method for increasing muscle mass in
 CC animals by administering an antibody (Ab) that binds to GDF-8; (3)
 CC inhibiting the action of GDF-8 by treating foetal or adult muscle
 CC or progenitor cells with a GDF-8 inhibitor; (4) isolated nucleic
 CC acid encoding a GDF-8 protein truncated by loss of the C-terminal
 CC active fragment. The transgenic animals have increased muscle mass
 CC and for poultry reduced cholesterol contents. Method (3) is used
 CC to treat muscle wasting or neuromuscular diseases, muscular atrophy
 CC and aging, particularly muscular dystrophy, spinal cord or
 CC traumatic injuries, congestive or obstructive lung disease, AIDS
 CC and cachexia. Method (4) is used to treat cancer of muscle,
 CC connective tissue and bone, or obesity. Also (not claimed) GDF-8
 CC can be used to maintain myoblasts intended for transplanting or to
 CC improve efficiency of fusion. Ab can be used to detect and
 CC quantify GDF-8 (particularly in muscle, for diagnosis or monitoring),
 CC also for immunotherapy and in vivo imaging.
 CC Sequence 375 AA;

W69887 Length: 375 February 11, 2000 15:49 Type: P Check: 9305 ..

```

1 MOKIOLSVYI YLFMLIVAGP VDLNENSEQK ENVEKEGICN ACLWRENTS
51 SRLFAIKIOI LSKRLLETAP NISKDAIROL LPKAPPLLEL IDQFDVORDA
101 SSDGSLDDDD YHATETVIT MPTESDLITQ VEGKPKCCF KESSKIQYKN
151 LYKAOLMIYL RPYETPATVF VOILRLIKPM KDGRTYTGIR SLKIDNPGT
201 GIWOSIDVKT VLOWMLKQPE SNLGEIRKAL DENGHDLAVT FPPGEGDLT
251 PLEVKYVDT PKRSRDFGL DCDHSTESR CCRYPLTVDF EALGWDIITA
301 PKRYKANCYS GECEVFLQK YPHTILVHOA NPGSAGBPC TPTKMSPINM
351 LYFNKGEOII YGKIPAMVVD RCGCS

```

!!AA_SEQUENCE 1.0
 ID W80799 standard; Protein: 244 AA.
 AC W80799;
 DT 08-DEC-1998 (first entry)
 DE Rhodococcus nitrate hydratase gene fragment product.
 KW Rhodococcus; nitrate hydratase gene; expression vector; microbe;
 KW control factor; nitrate hydratase gene promoter.
 OS Rhodococcus erythropolis.
 PN J10248578-A.
 PD 22-SEP-1998.
 PE 05-MAR-1997; 065618.
 PR 05-MAR-1997; JP-065618.
 PA (NITT) NITTO CHEM IND CO LTD.
 DR WPI: 98-560730/48.
 DR N-PSDB: V65530.
 PT Expression vector - for Rhodococcus genus microbe
 PS Claim 2; Pages 5-6; 10p; Japanese.
 CC This represents a product of the nitrate hydratase gene fragment used in
 CC the construction of an expression vector for a Rhodococcus genus microbe.
 CC The expression vector contains a DNA region encoding a control factor
 CC having the ability of activating a nitrate hydratase gene promoter, a DNA region
 CC of the nitrate hydratase gene promoter activated by the above control factor, a
 CC DNA region producible in the cell of a Rhodococcus genus microbe, and a
 CC microbe. An exotic gene can be expressed constitutionally.
 CC Sequence 244 AA;

W80799 Length: 244 February 11, 2000 15:49 Type: P Check: 2018

1 MAGADVAHOG GINRRARILY VDDKEHVRIM VTMQLESENF DVAADAGDA
51 ALRQVTSAP DLMVLDLISLPG GKGGLVLIAT VRTDALPIV VLNARDETE
101 RIVALDDGAD DYIKPFSR ELAARIRAVL RRTTAEPPHE AAVGRFSDLE
151 IDTAREVRIL HGIPLEETK EFDLAVYMAA SPWQVFSRRR LLEWNRSSP
201 DMQDQVATE HVRIRKRIE EDPKPTILO TYAGAGYRD GERA

11AA-SEQUENCE 1.0
ID W80824 standard; Protein: 189 AA.

AC 24-DEC-1998 (first entry)
DE S. pneumoniae GFP binding protein.
DE Streptococcus pneumoniae protein; recombinant; gene expression; DNA chip;
KW virulence; antibody; infection; detection; treatment; hypothetical;
KW cell wall biosynthetic; external target; minimal gene set protein.
KW Streptococcus pneumoniae.
PN WO9826072-A1.
PN 18-JUN-1998.
PF 09-DEC-1997; US-036281.
PA (ELIL) LILLY & CO ELI.
PI Balitz RH, Burgett SG, Denoff BS, Hoskins JA, Jaskunas SR,
PI Mills BJ, Norris FH, Peery RB, Rostock PR, Rostock PR,
PI Skatrud PL, Smith MC, Solenberg PJ, Treadway PJ,
PI Young Bellido ML;
DR WPI: 98-348529/30.
DR N-PSDB: V65202.
DR Streptococcus pneumoniae nucleic acid sequences - used in DNA chips
PT for evaluating gene expression, and identification of virulence
PI genes

PS Claim 3: Pages 210-211; 333pp; English.
CC This sequence represents a Streptococcus pneumoniae GFP-binding
CC protein. The invention provides DNA sequences (V65201 to V65304) from
CC the Streptococcus pneumoniae genome and corresponding protein sequences
CC (W80605 to W80728). The protein sequences are classified as hypothetical,
CC cell wall biosynthetic, external target, or minimal gene set proteins. A
CC recombinant host containing a vector comprising any of the above nucleic
CC acids can be used for the recombinant expression of the proteins. The
CC invention also provides a DNA chip having arrayed on it at least 15 base
CC pair fragments of any one or more of these DNA sequences. The DNA chip can
CC be used methods for evaluating gene expression in S. pneumoniae and for
CC identifying virulence genes in S. pneumoniae. Antibodies that selectively
CC bind to the above proteins or peptide fragments can be used to treat
CC S. pneumoniae infection. The antibodies can also be used to detect
CC S. pneumoniae cells.
SQ Sequence 189 AA;

W80624 Length: 189 February 11, 2000 15:49 Type: P Check: 4504

1 MSASGRDPY EDYLAINKEI ESYNRLMER POIIVNKKD MPESSENEE
51 FFKKIAENYD EPEELPAIFP ISGLTKGGLA TLLDARAELL DKTPEFLIYD
101 ESDMEBEVY GFDEEKAFK ISRDDATWV LSGEKLKMLF NNTNDRDES
151 VMKFAQLRG MGVDEALRAR GAKGDLVRI GKPEFEFVD

11AA-SEQUENCE 1.0
ID W68494 standard; Protein: 368 AA.

AC W68494;
DE 07-JAN-1999 (first entry)
DE E2 papillomavirus protein encoded by plasmid PCGE2.18VONCO.
KW E2 protein; apoptosis; HPV; infection; papillomavirus-associated cancer;
KW cervical; virus-infected cell; p53 tumour repressor.
OS Papillomavirus.
PN W6832861-A1.
PN 30-JUL-1998.
PD 29-JAN-1998; F00169.

29-JAN-1997; FR-000964.
PA (INSP) INST PASTEUR.
PA (UYME-) UNIV MEXICO NACIONAL AUTONOMA.
PA Demeret C, Desaintes C, Goyat S, Thierry F, Yaniv M;
DR WPI: 98-427957/36.
DR N-PSDB: V60833.
DR Papilloma virus E2 protein or nucleic acid encoding it - useful for
PT treating and preventing cancer of the cervix
PT disclosue; Fig 12; 11pp; French.
PS The present sequence is encoded by an insert of plasmid PCGE2.18VONCO
CC encoding an E2 papillomavirus protein. Compounds derived from E2 are able
CC to induce apoptosis in cells that have integrated part of the human
CC papillomavirus (HPV) genome. The E2 protein and its derivatives the
CC vectors (including those expressing wild-type E2) and the corresponding
CC proteins or truncated E2 (E2TR) are all useful for treating or preventing
CC papillomavirus infection, particularly papillomavirus-associated cancers
CC (especially of the cervix uteri). The proteins, and sequences expressing
CC them, also induce apoptosis of virus-infected cells and increasing the
CC activity of the p53 tumour repressor.
SQ Sequence 368 AA;

W68494 Length: 368 February 11, 2000 15:49 Type: P Check: 9982

1 TMETPKETLS ERLSCYQDKI IDHYENDSKD IDSQIQWQL IRWENALFFA
51 AREHGICQLN HOVPAVYNIS KSKAKHAIEL QVALGILAS RYKTEDWTIQ
101 DTCEELMATE PTHCFKKGQ TVQYVFDGK DNQMTVAVD SVYMTDAGT
151 WDKATCVSH RGLIYKEG NTFFERKSE CEKYGNTGTM EVHGNVVID
201 CNDSCMSTSD DYSATOLYK QLOHPSRYS STVSVGTAT YGQTSATRP
251 GHGGLAEKOH CGPYNPLGA APTGNKKR KLCSGMTPI IHKGNRSL
301 KCLRRLRKH SDHYRDISST WHWTGANEK TGLIYTHYS ETQRKFLNT
351 VAIPDSVQIL VGYMTGMS

11AA-SEQUENCE 1.0
ID W82504 standard; Protein: 225 AA.

AC W82504;
DE 01-FEB-1999 (first entry)
DE Human EPRG1 protein #1.
DE EPRG1; EPO primary response gene 1; diagnosis; gene therapy; immunity;
KW disease; vaccine; inoculate; antibody; T cell; anaemia; polycytemia;
KW cancer; neutropenia; AIDS; diabetes; myelosuppression; allergy; asthma;
KW autoimmune disease; inflammatory disease; chromosome mapping; human.
PN Homo sapiens.
PN EP-87/038-82.
PD 11-NOV-1998; 303597.
PD 07-MAY-1998; US-071342.
PD 07-MAY-1998; US-045890.
PR 07-MAY-1997; US-045890.
PA (SMIR) SMITHKLINE BECKMAN CORP.
PA Dillon S, Lord K;
PI WPI: 98-570499/49.
DR N-PSDB: V69307.
DR New EPO primary response gene polypeptides and polynucleotides -
PT useful as diagnostic reagents and for prevention and treatment of
PT cancer and autoimmune and inflammatory diseases
PT disclosue; Page 19-20; 25pp; English.
PS This sequence represents a novel human EPO primary response gene 1
CC (EPRG1) polypeptide. EPRG1 polypeptides and polynucleotides are useful
CC for diagnosis of a disease or susceptibility to a disease by detecting
CC mutations in the EPRG1 gene using probes containing the EPRG1 nucleotide
CC sequence, or determining EPRG1 polypeptide or mRNA expression levels
CC EPRG1 polypeptides can be used to screen for agonists and antagonists
CC which bind the EPRG1 polypeptide by measuring resulting mRNA levels with
CC ELISA. These can be used in treatment to activate (agonist) or inhibit
CC (antagonist eg EPRG1 ligand, receptor or substrate) EPRG1 activity, in
CC addition to eg administration of antisense sequences to prevent
CC expression, or EPRG1 polypeptides to treat conditions associated with

CC a lack of EPRG1 protein. Gene therapy may also be used to affect
 CC endogenous EPRG1 polypeptide production. EPRG1 antibodies are useful for
 CC inducing an immune response to immunise and prevent diseases, and for
 CC isolating EPRG1 clones or purifying the polypeptides by affinity
 CC chromatography. EPRG1 polypeptides can be administered directly or as a
 CC vaccine to inoculate against disease by inducing an antibody and T-cell
 CC response. Diseases diagnosed, prevented or treated include anaemia, diabetes,
 CC polycythemia, cancer, neutropenia, AIDS, drug-induced anaemia, diabetes,
 CC myelosuppression, autoimmune diseases, rheumatoid arthritis and multiple
 CC sclerosis, and inflammatory diseases, including asthma and allergies. The
 CC EPRG1 polypeptide is also useful for mapping the gene to a chromosome,
 CC allowing gene inheritance to be studied through linkage analysis. The
 CC 3'-UTR segment of EPRG1 RNA may be useful to screen for agents which
 CC modulate RNA stability and turnover rate.
 SQ Sequence 225 AA:

W82504 Length: 225 February 11, 2000 15:49 Type: P Check: 447 ..

1 MYTHKFRPA GMSRPDTSL RAKTESKSE YQLVNNAVRK LDSEGFYWSA
 51 VTGGGANLIL SAEPAQTFLI ROSSDQRHF TLSVKTQSGT KNLRICCEG
 101 SFSLSDDPRS TOPVPRFDCV LKIVHHYMP PGADSPSPSP TEPSSVEPQ
 151 PSAQPLPGSP PRRAYIYSG GERKPLVLSR PLSSNVAITQ HLCRTVNGH
 201 LDSYKVTQL PGPIREFLDQ YDAPL

!!AA_SEQUENCE 1.0
 ID W82505 standard; Protein: 157 AA.
 AC W82505:

DT 01-FEB-1999 (first entry)

DE Human EPRG1 protein #2.

KW EPRG1; EPO primary response gene 1; diagnosis; gene therapy; immunity;
 KW disease; vaccine; inoculate; antibody; T cell; anaemia; polycythemia;
 KW cancer; neutropenia; AIDS; diabetes; myelosuppression; allergy; asthma;
 KW autoimmune disease; inflammatory disease; chromosome mapping; human.

OS Homo sapiens.

PN EP-877030-A2.

PD 11-NOV-1998.

PF 07-MAY-1998; 303597.

PR 01-MAY-1998; US-071342.

PR 07-MAY-1997; US-045890.

PA (SMK) SMITHKLINE BECKMAN CORP.

PT Dillon S, Lord K.

DR WPI: 98-570498/49.

DR N-PSDB: V69308.

PT New EPO primary response gene polypeptides and polynucleotides -

PT useful as diagnostic reagents and for prevention and treatment of

PT cancer and autoimmune and inflammatory diseases

PS Claim 1; Page 2; 25pp; English.

CC This sequence represents a novel human EPO primary response gene 1

CC (EPRG1) polypeptide. EPRG1 polypeptides and polynucleotides are useful

CC for diagnosing a disease or susceptibility to a disease by detecting

CC mutations in the EPRG1 gene using probes containing the EPRG1 nucleotide

CC sequence, or determining EPRG1 polypeptide or mRNA expression levels

CC EPRG1 polypeptides can be used to screen for agonists and antagonists

CC which bind the EPRG1 polypeptide by measuring resulting mRNA levels with

CC ELISA. These can be used in treatment to activate (agonist) or inhibit

CC (antagonist eg EPRG1 ligand, receptor or substrate) EPRG1 activity, in

CC addition to direct administration of antisense sequences to prevent

CC expression, or EPRG1 polypeptides to treat conditions associated with

CC a lack of EPRG1 protein. Gene therapy may also be used to affect

CC endogenous EPRG1 polypeptide production. EPRG1 antibodies are useful for

CC inducing an immune response to immunise and prevent diseases, and for

CC isolating EPRG1 clones or purifying the polypeptides by affinity

CC chromatography. EPRG1 polypeptides can be administered directly or as a

CC vaccine to inoculate against disease by inducing an antibody and T-cell

CC response. Diseases diagnosed, prevented or treated include anaemia, diabetes,

CC polycythemia, cancer, neutropenia, AIDS, drug-induced anaemia, diabetes,

CC myelosuppression, autoimmune diseases, rheumatoid arthritis and multiple

CC sclerosis, and inflammatory diseases, including asthma and allergies. The

CC EPRG1 polypeptide is also useful for mapping the gene to a chromosome,

CC allowing gene inheritance to be studied through linkage analysis. The
 CC 3'-UTR segment of EPRG1 RNA may be useful to screen for agents which
 CC modulate RNA stability and turnover rate.
 SQ Sequence 157 AA:

W82505 Length: 157 February 11, 2000 15:49 Type: P Check: 9756 ..

1 EFANSSQGRH FQLSVKTS GTKNLRICQE GGSFQLQSDP RKNQPVPRVD
 51 CVLKVRSG APAGAPSPS PRPEPSEVP EPPSQPLPG SPRRAYIY
 101 SGGEKIPVL SRPLSSNVAI LQHLCKRTVN GHLDSEYKT QLPPIREFL
 151 DQYDAPL

!!AA_SEQUENCE 1.0
 ID W81740 standard; Protein: 120 AA.
 AC W81740:

DT 27-JAN-1999 (first entry)

DE M. tuberculosis immunogenic polypeptide RDIF5.

KW Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;
 KW vaccine; pharmaceutical; infection; diagnosis.

OS Mycobacterium tuberculosis.

PN W09816646-A2.

PD 23-APR-1998.

PF 07-OCT-1997; U18293.

PR 13-MAR-1997; US-818112.

PR 11-OCT-1996; US-730510.

PA (CORI-) CORIXA CORP.

PI Campos-Neto A, Dillon DC, Houghton R, Lodes MJ,

PI Reed SG, Skeiky YAM, Twardzik DR, Vedvick TS;

DR WPI: 98-261042/23.

DR N-PSDB: V64551.

PT Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used

PT to develop products for the detection of M. tuberculosis infection

PT and for diagnosis, treatment and prevention of tuberculosis

PS Example 3d; Page 190-191; 230pp; English.

CC This sequence represents an immunogenic portion of a soluble

CC Mycobacterium tuberculosis (MT) antigen which can be used in a method

CC for inducing protective immunity against tuberculosis (TB). This sequence

CC can be formulated into vaccines and/or pharmaceutical compositions for

CC immunising against M. tuberculosis infection or may be used for the

CC diagnosis of tuberculosis.

SQ Sequence 120 AA:

W81740 Length: 120 February 11, 2000 15:49 Type: P Check: 3575 ..

1 TPKEFYDDL DIDSLSWEI AVOTEKYG KIPDEDLACL RTVDGVYAI
 51 QKLEENPEA AQALRAKIES ENPDARADR CVSPSQARD ARPLANSAR
 101 LACRRLPASY PTRRDRER

!!AA_SEQUENCE 1.0

ID W72902 standard; Protein: 226 AA.

AC W72902:

DT 21-JAN-1999 (first entry)

DE Mycobacterium tuberculosis antigen CP19A.

KW Mycobacterium tuberculosis; antigen; vaccine; immunological;

OS Mycobacterium tuberculosis.

PN Key

FT Location/Qualifiers

FT Peptide

FT 1..38

FT /label= signal

PN W09844119-A1.

PD 08-OCT-1998.

PF 01-APR-1998; DK0132.

PR 05-JAN-1998; US-070488.

PR 02-APR-1997; DK-000376.

PR 18-APR-1997; US-044624.

PR 10-NOV-1997; DK-001277.

PA (STAT-) STAVENS SERUM INST.

PI Andersen P, Florio W, Nielsen R, Oettinger T, Rasmussen PB,

PI Rosenkrantz I, Weidling K:
DR N-PSDB; V63926.
PT New isolated mycobacteria polyepitopes and nucleic acids - used for
PT developing products for the diagnosis of or vaccination against
PT mycobacterial infections, particularly tuberculosis
PS Claim 1; Page 139; 10pp; English. Mycobacterium tuberculosis protein.
CC The present sequence represents a mycobacterium tuberculosis protein.
CC Products from the present invention which describes protein fragments
CC and nucleic acid fragments derived from M. tuberculosis, can be used in
CC the detection of and prevention of mycobacterial infections. In
CC particular, the proteins and nucleic acids can be used for the diagnosis
CC of or vaccination against tuberculosis caused by M. tuberculosis,
CC M. africanum or M. bovis.
SQ Sequence 226 AA;
WT2902 Length: 226 February 11, 2000 15:49 Type: P Check: 5027 ..
1 MIRPPOPHSG KWRAGARRL TSLVAAFAA ATLLTPALA PPASAGCPDA
51 EYVFAKTCG PPGICRGVCA FVSSLRQTN KSIGTYVNY PANGDFLAAA
101 DGANDADHI QGMASACRAI RLVIQGSOG AAVIDIYAA PLPLGFTCP
151 LPPADADHIA AIAIFGNPSG RAGGLMSALT PFGSKTINL CNNQDFICSD
201 GNRRAHLG VPGKNOAR FVASRI
11AA SEQUENCE 1.0
ID W80495 standard; protein; 207 AA.
AC W80495; 29-FEB-1999 (first entry)
DE Human vascular endothelial growth factor (VEGF)-B186.
DE Vascular endothelial growth factor; VEGF; proliferation; human;
KW endothelial cell; angiogenesis; tissue growth; organ repair.
KM Homo sapiens.
CN US5840693.A.
PN 24-NOV-1998.
PP 01-MAR-1996; 609443.
PP 01-MAR-1996; US-609443.
PP 01-MAR-1995; US-397651.
PP 06-JUN-1995; US-469427.
PP 06-DEC-1995; US-569063.
PA (LUDWIG) LUDWIG INST CANCER RES.
PA (UYHE-) UNIV HELSINKI LICENSING LTD OY.
PA Alitalo K, Eriksson U, Olofsson B, Pajusola K;
PI WPI: 99-034079/03.
DR N-PSDB; V63569.
PT Vascular endothelial growth factor-B isoforms, and DNA encoding
PT them - useful for inducing angiogenesis and cellular proliferation,
PT and raising antibodies to inhibit activities in e.g. tumours
PS Claim 1; Fig 15; 52pp; English.
CC The present sequence represents human vascular endothelial growth factor
CC (VEGF)-B186. VEGF proteins are used for promoting proliferation of
CC endothelial cells and for stimulating angiogenesis (the proliferation
CC of new capillaries from pre-existing blood vessels). These activities
CC are useful for treating tissue growth and repair, including organ
CC repair. This is also useful in pregnancy, in follicle development,
CC as these processes must occur in development of the placenta. The
CC proteins can also be used to raise antibodies, either for use in
CC detection of the proteins or as inhibitors of their action. This is
CC especially useful as angiogenesis is required by tumours as they need
CC new blood supplies to grow and proliferate.
SQ Sequence 207 AA;
W80495 Length: 207 February 11, 2000 15:49 Type: P Check: 1679 ..
1 MSPILRLRL ALLQIAPAQ APVSODAPD HOKRVSNID VYTRATCOPR
51 EYVPLPLVEL MGTVAQOLVP SCVYORCGG CCPDGLCEV PTGHOVORMQ
101 ILMIRPSSQ LGEMSLIEHS QCECRPKKD SAKKPDRAAT PHRPOPRSY
11AA SEQUENCE 1.0
ID W85306 standard; peptide; 24 AA.
AC W85306; 16-FEB-1999 (first entry)
DE Helper T-cell class II peptide derived from mouse invariant chain.
DE Helper T-cell peptide; human leucocyte antigen, HLA; DR4v; DR1;
KW DR7; cytotoxic T lymphocyte; CTL; hepatitis; autoimmune disease;
KW acquired immune deficiency syndrome; malaria; cancer
KW allograft rejection; allergy; Lyme disease; hepatitis;
KW post-streptococcal endocarditis; glomerulonephritis;
KW food hypersensitivity.
OS Synthetic.
PN W85306; 30-JUL-1998.
PD 23-FEB-1998; 001373.
PE 07-FEB-1997; US-057452.
PR 23-JAN-1997; US-056713.
PA (EPIM-) EPIMUNE, INC.
PI Sette A, Sidney J, Southwood S;
DR WPI: 98-427679/36.
PT Composition containing peptide that induces cytotoxic T lymphocyte
PT response, and helper peptide - can bind to human leucocyte antigen
PT alleles, used to treat or prevent cancers, parasitic infections and
PT autoimmune disease
PS Disclosure; Page 40; 51pp; English.
CC W85284-451 represent helper T-cell class II peptides, which can bind to
CC the human leucocyte antigens (HLA) DR4v, DR1 and DR7. The peptides
CC are used in the course of the invention, the specification describes
CC peptides that that induce a cytotoxic T lymphocyte (CTL) response, and
CC T-helper peptides, that are used together to generate a CTL response for
CC the treatment or prevention of viral, fungal, bacterial or parasitic
CC infections (e.g. hepatitis acquired immune deficiency syndrome or
CC malaria) or cancer (e.g. internal or cervical carcinoma, lymphoma, prostate
CC cancer or condyloma). Helper T-cell peptides may be used
CC alone to induce an allograft rejection, allergy, Lyme disease, hepatitis,
CC disease, allograft rejection, allergy, Lyme disease, hepatitis,
CC post-streptococcal endocarditis, glomerulonephritis and food
CC hypersensitivity.
SQ Sequence 24 AA;
W85306 Length: 24 February 11, 2000 15:49 Type: P Check: 3126 ..
1 LPKPKPVSK MRKATPLMQ ALPM
11AA SEQUENCE 1.0
ID W86202 standard; protein; 206 AA.
AC W86202; 16-FEB-1999 (first entry)
DE Human VEGF-related factor (VRF)-2 sequence.
DE VEGF; VRF; vascular endothelial growth factor; VEGF-related protein;
KW recombinant; truncated; gene therapy; angiogenesis; cardiac ischaemia;
KW lower limb ischaemia; stroke; peripheral vascular disease; intestine;
KW wound healing; skin; VEGF-related factor; VRF; vascular permeability.
KM Homo sapiens.
CN WO9849300-A2.
PN 05-NOV-1998.
PP 25-APR-1997; US-842984.
PA (COLL-) COLLATERAL THERAPEUTICS.
PI Bohlen P;
DR WPI: 99-009426/01.
PT New truncated vascular endothelial growth factor-related protein
PT subunits - lack part of the N-terminal sequence, used to stimulate
PT angiogenesis; e.g. for treating heart disease and ischaemia
PS Disclosure; Fig 1; 11pp; English.
CC This represents the amino acid sequence of human VRF-2 (a vascular
CC endothelial growth factor (VEGF)-related factor). The invention


```
CC Provides truncated VRP (VEGF-related protein) subunits that have at
CC least one amino acid N-terminal to the first Cys of the core sequence
CC deleted. Host cells transfected or transfected with expression vectors
CC containing nucleic acids encoding the truncated VRP subunits are used to
CC produce the truncated proteins recombinantly. The truncated VRP subunits,
CC optionally expressed from gene therapy vectors, have in vivo and in vitro
CC angiogenic activity and are used to stimulate angiogenesis, particularly
CC in the context of collateral vessel development in cases of cardiac ischemia; to
CC stimulate endothelial cell growth and migration in vitro; to treat heart
CC disease; to treat ischemia (e.g. cardiac, chronic coronary or chronic
CC lower limb ischemia), stroke and peripheral vascular disease); to promote
CC healing of wounds (of skin or intestines), and to increase vascular
CC permeability.
SO Sequence 206 AA;

W68202 Length: 206 February 11, 2000 15:49 Type: P Check: 9273
1 MSPLLRLL ALLQLAPQ AVSOPDAPG HQRKVSMD VYRATQPR
51 EVVPLIYEL MGVAKOLV SCVYVQRCGG CCPDGLCEV PTGQHQRNQ
101 LMIRYPSQ LGKMSLEHS QCECRPKDS AVKPDRAATP HHRPQRSVP
151 GWDAPGAPG PADITHPTA PGPSSAAPS TTSALTPGPA AAAADAASS
201 VAKGGA

11AA_SEQUENCE 1.0
ID W68541 standard; protein: 420 AA.
AC W68541:
DE 12-FEB-1999 (first entry)
DE Amino acid sequence of the STAX DP2-64 (Oct-T1) gene.
KW Tumour rejection antigen precursor: TRAP; TRH; leukemia; screening;
KW lymphoma; cancer; HLA; human lymphocyte antigen; vaccine;
KW STAX DP2-64; Oct-T1.
OS Homo sapiens.
PN MO9849299-A1.
PD 05-NOV-1998.
PF 23-APR-1998: U07784.
PR 25-APR-1997: US-845998.
PA (LUDW-) LUDWIG INST CANCER RES.
PI Boon-Fallert T, Coulle PG, De Smet C, Lucas S, Van Baren N;
DR N-PSDB: V33912.
PT New diagnosis of leukemia - by detecting genes for tumour antigen
PT rejection precursors or corresponding proteins
PS Claim 11: Pages 68-69: 88pp; English
CC The present sequence represents the amino acid sequence of the
CC STAX DP2-64 (Oct-T1) gene. The protein is a tumour rejection
CC antigen precursor (TRAP). The specification describes the treatment of
CC disorders which characterised by expression of a leukemia-associated
CC nucleic acid such as TRH. The products are used for in vivo or in vitro
CC screening for leukemia, lymphoma or other cancers by usual
CC hybridisation/amplification or immunosay methods. TRAPs, when
CC processed to antigens or complexed with HLA (human lymphocyte antigen)
CC molecules, or nucleic acid encoding them, are useful in vaccines for
CC treating leukemia.
SO Sequence 420 AA;

W68541 Length: 420 February 11, 2000 15:49 Type: P Check: 8848
1 MMSMSKQPH FAMHPTLPEH KYPSLSSE AIRRACLP PLOSNFASL
51 DETLLARAFA LAAYDIAVSQ GSKHPKPA TYHMINSPC TSTSTYPLAH
101 HHHHHHHQA LEPGDLDDH SSPSLALMAG AGGAGAAAG GGAHDOPGG
151 GCGGCGGGG GCGPGGGGG GCGGGGGGG GGLLGSNHP HPHMSLGH
201 SHPAAALMN MPGLPHGL VAAAHHGA AAAAAAAG VAAASAAAV
251 VGAAGLASIC DSDTPRELE AFAEFKQK IKLGTVQADV GSNLNLKIP

301 GVGSLSOSTI CREESLISH NMKILKPL QAWLENEGA OREKNNKPEL
351 FNGGKRRKR TSAAPKRS LEAYVAVOPR PSSERKIAIA EKLDLKNVY
401 KWFENQOK OKRKRKSATY

11AA_SEQUENCE 1.0
ID W68235 standard; protein: 185 AA.
AC W68235:
DE 16-FEB-1999 (first entry)
DE Human VRF (VEGF-related factor)-2 full length sequence.
KW VEGF; VRF; vascular endothelial growth factor; VEGF-related protein;
KW recombinant; truncated; gene therapy; angiogenesis; cardiac ischemia;
KW coronary; collateral vessel development; cell growth; migration; heart;
KW lower limb ischemia; stroke; peripheral vascular disease; intestine;
KW wound healing; skin; VEGF-related factor; VRF; vascular permeability.
OS Homo sapiens.
PN MO9849300-A2.
PD 05-NOV-1998.
PF 20-APR-1998: U07801.
PR 25-APR-1997: US-842984.
PA (COLL-) COLLATERAL THERAPEUTICS.
PI Bohlen P;
DR WPI: 99-009426/01.
PT New truncated vascular endothelial growth factor-related protein
PT subunits - lack part of the N-terminal sequence, used to stimulate
PT angiogenesis, e.g. for treating heart disease and ischemia
PS Claim 5; Fig 2b; 11pp; English
CC The invention relates to truncated VRP (vascular endothelial growth
CC factor (VEGF)-related protein) subunits that have at least one amino acid
CC N-terminal to the first Cys of the core sequence deleted. Host cells
CC transformed or transfected with expression vectors containing nucleic
CC acids encoding the truncated VRP subunits are used to produce the
CC truncated proteins recombinantly. The truncated VRP subunits, optionally
CC expressed from gene therapy vectors, have in vivo and in vitro angiogenic
CC activity and are used to stimulate angiogenesis, particularly coronary
CC collateral vessel development in cases of cardiac ischemia; to stimulate
CC endothelial cell growth and migration in vitro; to treat heart disease;
CC to treat ischemia (e.g. cardiac, chronic coronary or chronic lower limb
CC ischemia); stroke and peripheral vascular disease); to promote healing of
CC wounds (of skin or intestines), and to increase vascular permeability.
CC Sequences W68234 to W68239 represent full length VRF sequences from
CC which the truncated fragments are created.
SO Sequence 185 AA;

W68235 Length: 185 February 11, 2000 15:49 Type: P Check: 8227
1 PVSOPDAPG HQRKVSMD VYRATQPR VVPLIYEL MGVAKOLVS
51 CVTVQRCGG CCPDGLCEV TGOHQVRMGI LMIRYPSQ LGKMSLEHSQ
101 CECRPKDSA VKPDRAATP HHRPQRSVP WDSAPGAP ADITHPTAP
151 GPSAARST TSALTGPA AAAADAASSV AKGA

11AA_SEQUENCE 1.0
ID W68214 standard; protein: 178 AA.
AC W68214:
DE 16-FEB-1999 (first entry)
DE Human VRF-2 truncated fragment 1.
KW VEGF; VRF; vascular endothelial growth factor; VEGF-related protein;
KW recombinant; truncated; gene therapy; angiogenesis; cardiac ischemia;
KW coronary; collateral vessel development; cell growth; migration; heart;
KW lower limb ischemia; stroke; peripheral vascular disease; intestine;
KW wound healing; skin; VEGF-related factor; VRF; vascular permeability.
OS Homo sapiens.
PN MO9849300-A2.
PD 05-NOV-1998.
PF 20-APR-1998: U07801.
PR 25-APR-1997: US-842984.
PA (COLL-) COLLATERAL THERAPEUTICS.
PI Bohlen P;
DR WPI: 99-009426/01.
```


PI New truncated vascular endothelial growth factor-related protein
 PT subunits - lack part of the N-terminal sequence, used to stimulate
 PT angiogenesis, e.g. for treating heart disease and ischemia
 PS Claim 5; Fig 2B; 113pp; English.
 CC The invention relates to truncated VRF (vascular endothelial growth
 CC factor) (VEGF)-related protein subunits that have at least one amino acid
 CC N-terminal to the first Cys of the core sequence deleted. Host cells
 CC transformed or transfected with expression vectors containing nucleic
 CC acids encoding the truncated VRF subunits are used to produce the
 CC truncated proteins recombinantly. The truncated VRF subunits, optionally
 CC expressed from gene therapy vectors, have in vivo and in vitro angiogenic
 CC activity and are used to stimulate angiogenesis, particularly coronary
 CC collateral vessel development in cases of cardiac ischemia; to stimulate
 CC endothelial cell growth and migration in vitro; to treat heart disease;
 CC to treat ischemia (e.g. cardiac, chronic coronary or chronic lower limb
 CC ischemia); stroke and peripheral vascular disease); to promote healing of
 CC wounds (of skin or intestines), and to increase vascular permeability.
 CC Sequences W86214 to W86217 represent truncated fragments of VRF-2 (VEGF-
 CC related factor).
 SO Sequence 178 AA.

W86214 Length: 178 February 11, 2000 15:49 Type: P Check: 5797

1 PGHQRVSW IDVYTRATCQ PREVVVPLTV ELMGTVAKQL VPSCTVQRC
 51 GGCCEPDGILE CVPTGQHQVR MQLMIRYPS SOLGEMSLIE HSQCECRPK
 101 DSNVFPDRRA TPHRPQPRS VPGWDSAPGA PSPADITHPT PAPGSAHAA
 151 PPTTSALTGP PAAADAAA SSVAKGA

!!AA_SEQUENCE 1.0
 ID W86215 standard; protein: 173 AA.

AC W86215:
 DE 16-FEB-1999 (first entry)
 PT Human VRF-2 truncated fragment 2.
 CC VEGF: VRF: vascular endothelial growth factor; VEGF-related protein;
 CC recombinant; truncated; gene therapy; angiogenesis; cardiac ischemia;
 CC collateral vessel development; cell growth; migration; heart;
 CC lower limb ischemia; stroke; peripheral vascular disease; intestine;
 CC wound healing; skin; VEGF-related factor; VRF; vascular permeability.
 CC Homo sapiens.
 PN MO9849300-82.
 PD 05-NOV-1998.
 PF 20-APR-1998; U07801.
 PR 25-APR-1997; US-842984.
 PA (COLL-) COLLATERAL THERAPEUTICS.
 PI Bohlen P;
 PI WPI: 99-009426/01.
 DR New truncated vascular endothelial growth factor-related protein
 PT subunits - lack part of the N-terminal sequence, used to stimulate
 PT angiogenesis, e.g. for treating heart disease and ischemia
 PS Claim 5; Fig 2B; 113pp; English.
 CC The invention relates to truncated VRF (vascular endothelial growth
 CC factor) (VEGF)-related protein subunits that have at least one amino acid
 CC N-terminal to the first Cys of the core sequence deleted. Host cells
 CC transformed or transfected with expression vectors containing nucleic
 CC acids encoding the truncated VRF subunits are used to produce the
 CC truncated proteins recombinantly. The truncated VRF subunits, optionally
 CC expressed from gene therapy vectors, have in vivo and in vitro angiogenic
 CC activity and are used to stimulate angiogenesis, particularly coronary
 CC collateral vessel development in cases of cardiac ischemia; to stimulate
 CC endothelial cell growth and migration in vitro; to treat heart disease;
 CC to treat ischemia (e.g. cardiac, chronic coronary or chronic lower limb
 CC ischemia); stroke and peripheral vascular disease); to promote healing of
 CC wounds (of skin or intestines), and to increase vascular permeability.
 CC Sequences W86214 to W86217 represent truncated fragments of VRF-2 (VEGF-
 CC related factor).
 SO Sequence 173 AA.

W86215 Length: 173 February 11, 2000 15:49 Type: P Check: 4363
 1 KYSWIDVYT RATCPREVY VPLTVELMGT VAKQIVPSCV TVQRCGCCP

51 DDGLECVPTG QHQVQMQLM IRPSSOLGE MSLEHSCOE CRPKDSAYK
 101 PDRNATPHHR POPRSVPQWD SAGAPSPAD ITHPTAPGP SAHAPSTTS
 151 ALTPGPAAA ADAASVAK GGA

!!AA_SEQUENCE 1.0
 ID W86216 standard; protein: 168 AA.

AC W86216:
 DE 16-FEB-1999 (first entry)
 PT Human VRF-2 truncated fragment 3.
 CC VEGF: VRF: vascular endothelial growth factor; VEGF-related protein;
 CC recombinant; truncated; gene therapy; angiogenesis; cardiac ischemia;
 CC collateral vessel development; cell growth; migration; heart;
 CC lower limb ischemia; stroke; peripheral vascular disease; intestine;
 CC wound healing; skin; VEGF-related factor; VRF; vascular permeability.
 CC Homo sapiens.
 PN MO9849300-82.
 PD 05-NOV-1998.
 PF 20-APR-1998; U07801.
 PR 25-APR-1997; US-842984.
 PA (COLL-) COLLATERAL THERAPEUTICS.
 PI Bohlen P;
 PI WPI: 99-009426/01.
 DR New truncated vascular endothelial growth factor-related protein
 PT subunits - lack part of the N-terminal sequence, used to stimulate
 PT angiogenesis, e.g. for treating heart disease and ischemia
 PS Claim 5; Fig 2B; 113pp; English.
 CC The invention relates to truncated VRF (vascular endothelial growth
 CC factor) (VEGF)-related protein subunits that have at least one amino acid
 CC N-terminal to the first Cys of the core sequence deleted. Host cells
 CC transformed or transfected with expression vectors containing nucleic
 CC acids encoding the truncated VRF subunits are used to produce the
 CC truncated proteins recombinantly. The truncated VRF subunits, optionally
 CC expressed from gene therapy vectors, have in vivo and in vitro angiogenic
 CC activity and are used to stimulate angiogenesis, particularly coronary
 CC collateral vessel development in cases of cardiac ischemia; to stimulate
 CC endothelial cell growth and migration in vitro; to treat heart disease;
 CC to treat ischemia (e.g. cardiac, chronic coronary or chronic lower limb
 CC ischemia); stroke and peripheral vascular disease); to promote healing of
 CC wounds (of skin or intestines), and to increase vascular permeability.
 CC Sequences W86214 to W86217 represent truncated fragments of VRF-2 (VEGF-
 CC related factor).
 SO Sequence 168 AA.

W86216 Length: 168 February 11, 2000 15:49 Type: P Check: 2203

1 IDVYTRATCQ PREVVVPLTV ELMGTVAKQL VPSCTVQRC GGCCEPDGILE
 51 CVPTGQHQVR MQLMIRYPS SOLGEMSLIE HSQCECRPK DSNVFPDRRA
 101 TPHRPQPRS VPGWDSAPGA PSPADITHPT PAPGSAHAA PPTTSALTGP
 151 PAAADAAA SSVAKGA

!!AA_SEQUENCE 1.0
 ID W86217 standard; protein: 163 AA.

AC W86217:
 DE 16-FEB-1999 (first entry)
 PT Human VRF-2 truncated fragment 4.
 CC VEGF: VRF: vascular endothelial growth factor; VEGF-related protein;
 CC recombinant; truncated; gene therapy; angiogenesis; cardiac ischemia;
 CC collateral vessel development; cell growth; migration; heart;
 CC lower limb ischemia; stroke; peripheral vascular disease; intestine;
 CC wound healing; skin; VEGF-related factor; VRF; vascular permeability.
 CC Homo sapiens.
 PN MO9849300-82.
 PD 05-NOV-1998.
 PF 20-APR-1998; U07801.
 PR 25-APR-1997; US-842984.
 PA (COLL-) COLLATERAL THERAPEUTICS.
 PI Bohlen P;

DR WPI: 99-009426/01.
PT New truncated vascular endothelial growth factor-related protein
PT subunits - lack part of the N-terminal sequence, used to stimulate
PT angiogenesis, e.g. for treating heart disease and ischaemia
PS Clam 5; Fig 2B; 113pp; English.
CC The invention relates to truncated VRP (vascular endothelial growth
CC factor (VEGF)-related protein) subunits that have at least one amino acid
CC N-terminal to the first Cys of the core sequence deleted. Host cells
CC transformed or transfected with expression vectors containing nucleic
CC acids encoding the truncated VRP subunits are used to produce the
CC truncated proteins recombinantly. The truncated VRP subunits, optionally
CC expressed from gene therapy vectors, have in vivo and in vitro angiogenic
CC activity and are used to stimulate angiogenesis, particularly coronary
CC collateral vessel development in cases of cardiac ischaemia; to stimulate
CC endothelial cell growth and migration in vitro; to treat heart disease;
CC to treat ischaemia (e.g. cardiac chronic coronary or chronic lower limb
CC ischaemia, stroke and peripheral vascular disease); to promote healing of
CC wounds (of skin or intestines), and to increase vascular permeability.
CC Sequences W86214 to W86217 represent truncated fragments of VRF-2 (VEGF-
CC related factor).
SQ Sequence 163 AA;

W86217 Length: 163 February 11, 2000 15:49 Type: P Check: 3123 ..

1 RATCOPREVV VPLTVELMGT VAKOLVPCV TVORCGCCP DDGLECVPTG
51 OHQVROMILM IRYPSQIGE MSLEHSGCE CRRKDSAVK PDRAATPHNR
101 POPRSVPQWD SAPGAPSPAD ITHPTAPGP SAHAAPTST ALTPGMAAA
151 ADAASVYAK GCA

!!AA_SEQUENCE 1.0
ID W85119 standard; Protein: 446 AA.
AC W85119.
DE 11-FEB-1999 (first entry)
DE A delta-5 desaturase enzyme.
KW Fatty acid; delta-5 desaturase; polyunsaturated fatty acid;
KW malnutrition; inflammation; rheumatoid arthritis; asthma; psoriasis;
KW cancer; diabetes; eczema; platelet aggregation; vasodilation;
KW cholesterol level; endometritis; premenstrual syndrome;
KW myalgic encephalomyelitis; chronic fatigue; AIDS; multiple sclerosis;
KW acute respiratory syndrome; hypertension; inflammatory skin disorder.
OS Mortierella alpina.
PN WO9646765-A1.
PD 22-OCT-1998.
PF 10-APR-1998; U07422.
PR 11-APR-1997; US-833610.
PA (ABBO) ABBOTT LAB.
PA (CALJ) CALGENE LLC.
PI Chaudhary S, Huang Y, Knutson D, Leonard AE, Mukerji P,
PI Thurmond J.
DR WPI: 99-009334/01.
DR N-PSDB: W82628.
PT New nucleic acid encoding delta5 and other desaturase enzymes -
PT useful in production of oils of increased arachidonic acid content,
PT used, e.g. for treating cancer, as foods, animal feeds and cosmetics
PS Clam 3; Fig 3A-D; 153pp; English.
CC The present sequence represents a Mortierella alpina fatty acid delta-5
CC desaturase enzyme. The enzyme sequence is used in the methods of
CC the invention. The specification describes methods for desaturating a
CC fatty acid and for producing a desaturated fatty acid by expressing
CC increased levels of a desaturase. The present desaturase is an enzyme
CC which introduces a double bond carbons 5 and 6 from the carboxyl end of
CC a fatty acid molecule. The enzyme can be used for desaturating fatty
CC acids. The enzyme can be used to produce polyunsaturated fatty acids,
CC which can be used for treating malnutrition, in pharmaceutical
CC compositions, in cosmetics or in animal feed. The polyunsaturated fatty
CC acids can be used for treating e.g. restenosis after angioplasty,
CC inflammation, rheumatoid arthritis, asthma, psoriasis, cancer, diabetes
CC or eczema or reduce blood pressure. They can also be used to inhibit
CC platelet aggregation, cause vasodilation, lower cholesterol levels,
CC inhibit proliferation of vessel wall smooth muscle and fibrous tissue,

CC reduce or prevent gastro-intestinal bleeding and other side effects
CC caused by non-steroidal anti-inflammatory drugs, prevent or treat
CC endometritis and premenstrual syndrome, treat myalgic encephalomyelitis
CC and chronic fatigue after viral infections, treat AIDS, multiple
CC sclerosis, acute respiratory syndrome, hypertension and inflammatory skin
CC disorders.
SQ Sequence 446 AA;

W85119 Length: 446 February 11, 2000 15:49 Type: P Check: 5663 ..

1 MCTDQKXFT WEELAHNTR DOLLAIKGR VYDTKFLSR HPGVDITLL
51 GAGROVTVEF EMYHAFGAD AIMKKYVGT LVSNELDIFP EPIVFKITIK
101 TVEGEYFDR NIDPKNPEI WGRVALIFGS LIASYNOLF VPPVVERTML
151 QVVEALINGF ACAQVGNPL HDASHFSVTH NPTWKILGA THDFENGASY
201 LVWATQHNLF HHPYTNAGA DDVSTSEPD VRRIRPKQKW FVNHINQMF
251 VPELYGLIAF KVRIDINIL YFVKNDAIR VNPISWHVY MFWGKAFFV
301 WYRLIVPLQY LPKGVLLF TVADWSSYW LALTFQANH VEEVQMLPD
351 ENGIIOKDMA AMOVETQDY AHDSHLWTSI TGSILNYQAVH HLEPNVSOHH
401 YPDILAIKRN TCSEYKVPYL VKDFEWOAFA SHLEHLVLC LRPKEE

!!AA_SEQUENCE 1.0
ID W73507 standard; Protein: 317 AA.
AC W73507.
DE 01-MAR-1999 (first entry)
DE Human ATG-1709 protein.
KW Human; ATG-1709 protein; secreted ligand; 7-Transmembrane receptor;
KW heart disease; hypertension; cardiovascular disease; kidney disease;
KW obesity; insulin resistance; diabetes; Central Nervous System disorder;
KW therapy; sFRP-1.
OS Homo sapiens.
PN EP-879685-A1.
PD 23-NOV-1998.
PF 16-JAN-1998; 300313.
PR 16-AUG-1997; US-907808.
PR 23-MAY-1997; US-047691.
PA (SMIX) SMITHKLINE BEECHAM CORP.
PI Hu E, Zhu Y.
PI Hu E, Zhu Y.
DR WPI: 98-596877/51.
DR N-PSDB: V08946.
PT New human secreted protein ATG-1709 polypeptide and polynucleotide -
PT useful as diagnostic reagents and for diagnosing, prevention and
PT treatment of Central Nervous System diseases and diabetes
PS Clam 11; Page 6; 28pp; English.
CC This sequence represents the human ATG-1709 protein of the invention.
CC ATG-1709 is related to human secreted ligands for 7-Transmembrane
CC receptors and similar to murine sFRP-1. ATG-1709 polypeptides and
CC polynucleotides are useful for diagnosing susceptibility to diseases by
CC detecting mutations in the ATG-1709 gene using probes containing the
CC ATG-1709 nucleotide sequence, and can diagnose diseases associated with
CC ATG-1709 imbalance by determining ATG-1709 polypeptide expression levels.
CC ATG-1709 polypeptides can be used to screen for agonists and antagonists
CC which bind the ATG-1709 polypeptide. These can be used in treatment to
CC activate or inhibit ATG-1709 activity. In addition to direct
CC administration of antisense sequences to prevent expression, or ATG-1709
CC polypeptides to treat conditions associated with a lack of ATG-1709.
CC Gene therapy may also be used to affect endogenous ATG-1709 expression.
CC ATG-1709 antibodies are useful for inducing an immune response to
CC immunize and prevent diseases, and for isolating ATG-1709 clones or
CC purifying the polypeptides by affinity chromatography. ATG-1709
CC polypeptides can be administered directly or as a vaccine to inoculate
CC against disease. Diseases diagnosed, prevented or treated include:
CC heart disease; hypertension; cardiovascular diseases; kidney diseases;
CC obesity; insulin resistance; diabetes and Central Nervous System (CNS)
CC diseases. The ATG-1709 polypeptide is also useful for mapping the gene to
CC a chromosome, allowing gene inheritance to be studied through linkage

Mon Feb 14 08:07:18 2000

ags.cat

Page 87

CC analysis. 317 AA;
SQ Sequence Length: 317 February 11, 2000 15:49 Type: P Check: 3945
W73507
1 MRAAAGV RTAALLLLG ALHAPARCE EHYHYQMOE PLHGRYSKSP
51 PQCDIADRL PLCHTGYGR MRLPLLEHE SLAEVQOAS SWLPPLARC
101 HSDTVELCS LEAPVCLDR IYPCSLCEA VRAGCAPME AGCFWPEML
151 HCHRPDLND LCTAVOPGH PATAPVTKI CAQCEMHSA DGLMOMCSS
201 DFVAKRIKE IKIENDRKL IGAOKKKLL KPQELKDKT KLVYHMKNG
251 AGCEPQDLS LAGSFLVGR KYDGLLMA YRMKKKKE MKEAVKEMS
301 YPCSLYEPF YGAEPH
11AA_SEQUENCE 1.0
ID W90011 standard; Protein; 343 AA.
AC W90011;
DE 18-FEB-1999 (first entry)
KW Expressed antigen for cluster 35a.
KW Antigen: immunogenic cluster family; vaccine; gastritis; diagnosis;
KW Peptic ulcer; gastric adenocarcinoma; gastric lymphoma.
OS Helicobacter pylori.
PN MO9849314-A2.
PD 05-NOV-1998.
PF 27-APR-1998; U08487.
PR 14-OCT-1997; US-061958.
PR 25-APR-1997; US-045107.
PA (GENE-) GENELABS TECHNOLOGIES INC.
PI Chow TP, Fry KE, Lim MY, Mcatee CP;
DR WPI: 99-009433/01
PT New Helicobacter pylori antigens and related nucleic acid sequences
PT useful in serological diagnosis and protective vaccines, providing
PT long-lasting immune response
PS Claim 16; Page 338; 402pp; English.
CC The present sequence represents a Helicobacter pylori antigenic protein
CC that is characterised by immunoreactivity with H. pylori positive
CC antisera. The proteins are highly immunogenic and induce a long-lasting
CC immune response that persists even after antimicrobial treatment. In
CC antibody-detection assays, on sera, plasma, urine, saliva etc., they are
CC highly sensitive and specific. The specification also describes 69
CC previously unrecognised immunogenic cluster families. H. pylori antigens
CC are used to detect H. pylori-specific antibodies for diagnosis
CC infection or to confirm eradication of infection, and in vaccines to
CC protect against H. pylori infection and related diseases (gastritis,
CC peptic ulcer, gastric adenocarcinoma/lymphoma).
SQ Sequence 343 AA;
W90011 Length: 343 February 11, 2000 15:49 Type: P Check: 9038
1 MATKLPKOK AOLDELSMSE KIAILLIQG EDTGELIHR LIDISTEIS
51 KOIVOLNGTD KOIGAVILEE FFAIFOSNOY INTGLELEYAR ELLTRILGSE
101 EAKKVDKRLT KSLQOKNFA YLGKIRPOL ADFFINEHQ TIALIHAHE
151 APNAETLSY FPDEKAEIS IRMANGEIS POYKRVSTV LENKLESLNS
201 KYIEVGGLRA VAEIFNRLGQ KSAKTLIARI ESDVKRLAGA IKEMFTFED
251 IYKLDNFAR ELKRVADKRD LSLALTSRK DLTOKFLNM SSRAAEQVE
301 EMQYLGAVKI KDYDVAQRKI IEIVOSLOEK GVIQGEED VIE
11AA_SEQUENCE 1.0
ID W89910 standard; Protein; 343 AA.
AC W89910;
DE 18-FEB-1999 (first entry)
KW Antigen 1 from cluster 35a.

KW Antigen: immunogenic cluster family; vaccine; gastritis; diagnosis;
KW Peptic ulcer; gastric adenocarcinoma; gastric lymphoma.
OS Helicobacter pylori.
PN MO9849314-A2.
PD 05-NOV-1998.
PF 27-APR-1998; U08487.
PR 14-OCT-1997; US-061958.
PR 25-APR-1997; US-045107.
PA (GENE-) GENELABS TECHNOLOGIES INC.
PI Chow TP, Fry KE, Lim MY, Mcatee CP;
DR WPI: 99-009433/01
PT New Helicobacter pylori antigens and related nucleic acid sequences
PT useful in serological diagnosis and protective vaccines, providing
PT long-lasting immune response
PS Claim 1; Page 243-244; 402pp; English.
CC The present sequence represents a Helicobacter pylori antigenic protein
CC that is characterised by immunoreactivity with H. pylori-positive
CC antisera. The proteins are highly immunogenic and induce a long-lasting
CC immune response that persists even after antimicrobial treatment. In
CC antibody-detection assays, on sera, plasma, urine, saliva etc., they are
CC highly sensitive and specific. The specification also describes 69
CC previously unrecognised immunogenic cluster families. H. pylori antigens
CC are used to detect H. pylori-specific antibodies for diagnosis
CC infection or to confirm eradication of infection, and in vaccines to
CC protect against H. pylori infection and related diseases (gastritis,
CC peptic ulcer, gastric adenocarcinoma/lymphoma).
SQ Sequence 343 AA;
W89910 Length: 343 February 11, 2000 15:49 Type: P Check: 8829
1 MATKLPKOK AOLDELSMSE KIAILLIQG EDTGELIHR LIDISTEIS
51 KOIVOLNGTD KOIGAVILEE FFAIFOSNOY INTGLELEYAR ELLTRILGSE
101 EAKKVDKRLT KSLQOKNFA YLGKIRPOL ADFFINEHQ TIALIHAHE
151 APNAETLSY FPDEKAEIS IRMANGEIS POYKRVSTV LENKLESLNS
201 KYIEVGGLRA VAEIFNRLGQ KSAKTLIARI ESDVKRLAGA IKEMFTFED
251 IYKLDNFAR ELKRVADKRD LSLALTSRK DLTOKFLNM SSRAAEQVE
301 EMQYLGAVKI KDYDVAQRKI IEIVOSLOEK GVIQGEED VIE
11AA_SEQUENCE 1.0
ID W85390 standard; Protein; 354 AA.
AC W85390;
DE 02-MAR-1999 (first entry)
KW Polypeptide with 98% homology with human Wnt-11.
KW Wnt-11; disease; detection; cancer; cardiovascular disease; stroke;
KW Parkinsons disease; schizophrenia; alzheimers disease; acromegaly;
KW vitelliform macular dystrophy; vitreoretinopathy; assay;
KW neurological disease.
OS Homo sapiens.
PN EP-882791-A2.
PD 09-DEC-1998.
PF 20-MAY-1998; 303992.
PR 12-MAR-1998; GB-005332.
PR 23-MAY-1997; GB-010755.
PA (SMIR) SMITHKLINE BEECHAM PLC.
PI Barnes MR, Testa TT;
DR WPI: 99-011648/02.
DR N-PSDB: V83106.
PT New human Wnt-11 polypeptide - used to diagnose, treat and prevent
PT cancer, cardiovascular disease, developmental disorders, stroke and
PT neurological disease
PS Claim 1; Page 16-17; 21pp; English.
CC This polypeptide has at least 98 % sequence identity with the human
CC Wnt-11 345 amino acid sequence. Host cells are used to produce
CC this polypeptide from recombinant constructs which is then used
CC to screen for specific agonists and antagonists (potential therapeutic
CC agents) and to generate Ab (including induction of a protective immune
CC response). Human Wnt-11 polypeptide is involved in developmental

CC processes and its agonists and antagonists are used for treatment or
CC prevention of cancer, cardiovascular disease, stroke, neurological
CC disease (e.g. Parkinson's or Alzheimer's diseases, bipolar/unipolar
CC disorders and schizophrenia) or developmental disorders such as
CC acromegaly, vitelliform macular dystrophy or vitreoretinopathy.
CC Fragments of this polynucleotide sequence can be used for isolating
CC related genomic and cDNA clones and for chromosomal or tissue
CC localisation (the Wnt-11 gene is at chromosome 11q13). The specified
CC diseases, or susceptibility to them, are diagnosed by detecting a
CC mutation in a genomic Wnt-11 encoding sequence and/or by analysis for
CC presence or amount of Wnt-11 (or related mRNA) in a patient sample.
SQ Sequence 354 AA;

W85590 Length: 354 February 11, 2000 15:49 Type: P Check: 2306 ..

1 MRAPQCEA LLEFALQTG VCGIKWAL SKTSPALALN QTHCKQLES
51 LVSQVQLCR SNEIMHTV HAREVMKAC RRAFADMRN CSSIELAPN
101 LLDLRGTRE SAFYALSA AISHALRAC TSGDLPGSC GPYPGEPGP
151 GNRWGGCADN LSYGLMGAK FSDAPMKVK TGSQANKIMR LHNSEVGRQ
201 LRASLEMKCK CHGVSGSCI RTCKWGLQEL QDVADLCTR YLSATVHR
251 PMGTRKHLP KDLDIRPVK SELVYLQSSP DECKMKEKVG SHGTDRQCN
301 KTSNGSDSCD LMCGRGYNP YTDREVERCH CKYWCCTVT CRRCERTVER
351 YVCK

!!AA_SEQUENCE 1.0
ID W85591 standard; Protein; 351 AA.
AC W85591.

DE Polypeptide with 95% homology with human Wnt-11.
KW Wnt-11 disease; detection; cancer; cardiovascular disease; stroke;
KW parkinsons disease; schizophrenia; alzheimers disease; acromegaly;
KW vitelliform macular dystrophy; vitreoretinopathy; assay;
KW neurological disease.
OS Homo sapiens.
PN EP-882791-A2.
PD 09-DEC-1998.
PE 20-MAY-1998; 303992.
PR 12-MAR-1998; GB-005332.
PR 23-MAY-1997; GB-010755.
PA (SMIX) SMITHKLINE BEECHAM PLC.
PI Barnes MR, Testa TT;
DR WPI: 99-011648/02.
DR N-PDB: V83107.
PT New human Wnt-11 polypeptide - used to diagnose, treat and prevent
PT cancer, cardiovascular disease, developmental disorders, stroke and
PT neurological disease.
PS Claim 17: Page 18-19: 21pp: English.
CC This polypeptide has at least 95 % sequence identity with the human
CC Wnt-11 345 amino acid sequence. Host cells are used to produce
CC this polypeptide from recombinant constructs which is then used
CC to screen for specific agonists and antagonists (potential therapeutic
CC agents) and to generate Ab (including induction of a protective immune
CC response). Human Wnt-11 polypeptide is involved in developmental
CC processes and its agonists and antagonists are used for treatment or
CC prevention of cancer, cardiovascular disease, stroke, neurological
CC disease (e.g. Parkinson's or Alzheimer's diseases, bipolar/unipolar
CC disorders and schizophrenia) or developmental disorders such as
CC acromegaly, vitelliform macular dystrophy or vitreoretinopathy.
CC Fragments of this polynucleotide sequence can be used for isolating
CC related genomic and cDNA clones and for chromosomal or tissue
CC localisation (the Wnt-11 gene is at chromosome 11q13). The specified
CC diseases, or susceptibility to them, are diagnosed by detecting a
CC mutation in a genomic Wnt-11 encoding sequence and/or by analysis for
CC presence or amount of Wnt-11 (or related mRNA) in a patient sample.
SQ Sequence 351 AA;

W85591 Length: 351 February 11, 2000 15:49 Type: P Check: 8607 ..
1 RPQCEAMLF ALALQTCVY GIKWALSKT PSLALNQTH CKQLEGIVS
51 AQVOLCSNLT ELMHTVHAA REVMKACRRA FADMRNCS IELAPNLLVD
101 LERGTRESAF VYALSAAIS HAIRACTSG DLPGSCGPV GPEPPGPNR
151 WGGCADNLSTY WLIMGAKFSD APMKVKKTGS QANKLMGLIN SEVGRQALRA
201 SLEMKCKRGV VSGHCYIGTS WKGLQELQDV ADLKTTRYLS ATKVHRPMG
251 TRKHLVPKDL DIRPVKSEL VYLOSPDFC MKNEKGSNG TODRCNRTS
301 NGSDSCDLMC CGRGYNPTD RVERCHCKY HMCCTYTCRR CERTVERVVC
351 K

!!AA_SEQUENCE 1.0
ID W85100 standard; Protein; 36 AA.
AC W85100.

DE Polypeptide fragment encoded by gene 206.
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW endocrine; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
OS Homo sapiens.
PN WC9854963-A2.
PD 10-DEC-1998.
PE 04-JUN-1998; U11422.
PR 18-DEC-1997; US-048923.
PR 06-JUN-1997; US-048877.
PR 06-JUN-1997; US-048881.
PR 06-JUN-1997; US-048884.
PR 06-JUN-1997; US-048893.
PR 06-JUN-1997; US-048896.
PR 06-JUN-1997; US-048899.
PR 06-JUN-1997; US-048913.
PR 06-JUN-1997; US-048949.
PR 06-JUN-1997; US-048964.
PR 06-JUN-1997; US-048972.
PR 06-JUN-1997; US-049020.
PR 06-JUN-1997; US-049375.
PR 05-SEP-1997; US-057628.
PR 05-SEP-1997; US-057635.
PR 05-SEP-1997; US-057644.
PR 05-SEP-1997; US-057647.
PR 05-SEP-1997; US-057650.
PR 05-SEP-1997; US-057651.
PR 05-SEP-1997; US-057657.
PR 05-SEP-1997; US-057661.
PR 05-SEP-1997; US-057764.
PR 05-SEP-1997; US-057770.
PR 05-SEP-1997; US-057775.
PR 05-SEP-1997; US-057778.
PR 06-JUN-1997; US-048878.
PR 06-JUN-1997; US-048882.
PR 06-JUN-1997; US-048885.
PR 06-JUN-1997; US-048894.
PR 06-JUN-1997; US-048907.
PR 06-JUN-1997; US-048916.
PR 06-JUN-1997; US-048962.
PR 06-JUN-1997; US-048970.
PR 06-JUN-1997; US-048974.
PR 06-JUN-1997; US-049373.
PR 05-SEP-1997; US-057584.

PR 05-SEP-1997: US-057629.
 PR 05-SEP-1997: US-057642.
 PR 05-SEP-1997: US-057645.
 PR 05-SEP-1997: US-057648.
 PR 05-SEP-1997: US-057651.
 PR 05-SEP-1997: US-057652.
 PR 05-SEP-1997: US-057658.
 PR 05-SEP-1997: US-057752.
 PR 05-SEP-1997: US-057755.
 PR 05-SEP-1997: US-057771.
 PR 05-SEP-1997: US-057776.
 PR 06-JUN-1997: US-048876.
 PR 06-JUN-1997: US-048880.
 PR 06-JUN-1997: US-048883.
 PR 06-JUN-1997: US-048892.
 PR 06-JUN-1997: US-048895.
 PR 06-JUN-1997: US-048901.
 PR 06-JUN-1997: US-048910.
 PR 06-JUN-1997: US-048917.
 PR 06-JUN-1997: US-048963.
 PR 06-JUN-1997: US-048971.
 PR 06-JUN-1997: US-049019.
 PR 06-JUN-1997: US-049374.
 PR 05-SEP-1997: US-057627.
 PR 05-SEP-1997: US-057634.
 PR 05-SEP-1997: US-057643.
 PR 05-SEP-1997: US-057646.
 PR 05-SEP-1997: US-057654.
 PR 05-SEP-1997: US-057656.
 PR 05-SEP-1997: US-057659.
 PR 05-SEP-1997: US-057763.
 PR 05-SEP-1997: US-057765.
 PR 05-SEP-1997: US-057769.
 PR 05-SEP-1997: US-057774.
 PR 05-SEP-1997: US-057777.
 PR 05-SEP-1997: US-057777.
 PR (HUMAN) HUMAN GENOME SCI INC.
 PI Brewer LA, Carter KC, Dillon PJ, Ebner R, Endress GA,
 PI Fan P, Feng P, Ferrie AM, Fischer CT, Florence C,
 PI Florence K, Greene JM, Hu J, Kyaw H, Lafleur DM,
 PI Li Y, Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM,
 PI Shi Y, Soppet DR, Wei Y, Young P, Yu G, Zeng Z;
 PI WPI: 99-059865/05.
 DR N-PSDB: V84616.
 PT New isolated human genes and the secreted polypeptides they encode -
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders
 PS Disclosure: Page 159: 712pp; English. sequences (V84411 to V84633)
 CC The invention relates to nucleic acid sequences (V84411 to V84633)
 CC encoding human secreted proteins (W85354 to W88756). The secreted protein
 CC gene sequences are deposited with the ATCC under deposit numbers ATCC
 CC 97979, 97974, 97975, 97976, 97977, 209007, 209008, 209009, 209010, Host
 CC 209011, 209080, 209081, 209082, 209083, 209084, 209085, 209511. Host
 CC cells comprising recombinant vectors containing the nucleic acid
 CC sequences are used for the recombinant production of the secreted
 CC proteins. The polynucleotide and amino acid sequences are useful for are
 CC useful for preventing, treating or ameliorating medical conditions e.g.
 CC by protein or gene therapy. Pathological conditions can be also diagnosed
 CC by determining the amount of the new polypeptides in a sample or by
 CC determining the presence of mutations in the new polynucleotides, based on
 CC specific uses are described for each of the polynucleotides, based on
 CC which tissues they are most highly expressed in, and include developing
 CC products for the diagnosis or treatment of cancer, neurodegenerative
 CC disorders, developmental abnormalities and foetal deficiencies, blood
 CC disorders, tumours, leukemias, diseases of the immune system, autoimmune
 CC diseases, hepatic and renal disease, lymphomas, inflammation, allergies,
 CC ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,
 CC restenosis, prostate diseases, obesity, disorders involving osteoclasts
 CC such as osteoporosis, arthritis or malignancies, diseases of testes, lung
 CC or thymus, digestive/endocrine disorders, infections and AIDS. The
 CC polypeptides are also useful for identifying their binding partners.
 CC The present sequence represents a polypeptide fragment encoded by a
 CC gene of the invention (see descriptor line for gene number).
 SQ Sequence 36 AA:

W89100 Length: 36 February 11, 2000 15:49 Type: P Check: 1342 ...
 1 FSVHPELLE NISRFLLHSL PKDTPSGISK VILFT
 ID W88568 standard: Protein; 221 AA.
 AC W88568:
 DI 01-MAR-1999 (first entry)
 DE Secreted protein encoded by gene 35 clone HMDA574.
 KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;
 KW immune system; ischemic shock; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 OS Homo sapiens.
 PN W09854963-A2.
 PD 10-DEC-1998.
 PE 18-DEC-1997: US-070923.
 PR 06-JUN-1997: US-048877.
 PR 06-JUN-1997: US-048881.
 PR 06-JUN-1997: US-048884.
 PR 06-JUN-1997: US-048893.
 PR 06-JUN-1997: US-048896.
 PR 06-JUN-1997: US-048915.
 PR 06-JUN-1997: US-048919.
 PR 06-JUN-1997: US-048964.
 PR 06-JUN-1997: US-048972.
 PR 06-JUN-1997: US-049020.
 PR 05-SEP-1997: US-057628.
 PR 05-SEP-1997: US-057635.
 PR 05-SEP-1997: US-057644.
 PR 05-SEP-1997: US-057647.
 PR 05-SEP-1997: US-057650.
 PR 05-SEP-1997: US-057651.
 PR 05-SEP-1997: US-057661.
 PR 05-SEP-1997: US-057764.
 PR 05-SEP-1997: US-057766.
 PR 05-SEP-1997: US-057770.
 PR 05-SEP-1997: US-057775.
 PR 05-SEP-1997: US-057778.
 PR 06-JUN-1997: US-048875.
 PR 06-JUN-1997: US-048878.
 PR 06-JUN-1997: US-048882.
 PR 06-JUN-1997: US-048885.
 PR 06-JUN-1997: US-048894.
 PR 06-JUN-1997: US-048897.
 PR 06-JUN-1997: US-048900.
 PR 06-JUN-1997: US-048916.
 PR 06-JUN-1997: US-048962.
 PR 06-JUN-1997: US-048970.
 PR 06-JUN-1997: US-048974.
 PR 06-JUN-1997: US-049373.
 PR 05-SEP-1997: US-057584.
 PR 05-SEP-1997: US-057629.
 PR 05-SEP-1997: US-057642.
 PR 05-SEP-1997: US-057645.
 PR 05-SEP-1997: US-057651.
 PR 05-SEP-1997: US-057652.
 PR 05-SEP-1997: US-057658.
 PR 05-SEP-1997: US-057762.
 PR 05-SEP-1997: US-057765.
 PR 05-SEP-1997: US-057771.
 PR 05-SEP-1997: US-057776.
 PR 06-JUN-1997: US-048876.
 PR 06-JUN-1997: US-048880.

Mon Feb 14 08:07:18 2000

agus.cat

Page 90

06-JUN-1997: US-048883.
06-JUN-1997: US-048892.
06-JUN-1997: US-048895.
06-JUN-1997: US-048898.
06-JUN-1997: US-048901.
06-JUN-1997: US-048917.
06-JUN-1997: US-048963.
06-JUN-1997: US-048971.
06-JUN-1997: US-049019.
06-JUN-1997: US-049374.
06-JUN-1997: US-057627.
05-SEP-1997: US-057634.
05-SEP-1997: US-057646.
05-SEP-1997: US-057649.
05-SEP-1997: US-057654.
05-SEP-1997: US-057666.
05-SEP-1997: US-057760.
05-SEP-1997: US-057763.
05-SEP-1997: US-057769.
05-SEP-1997: US-057774.
05-SEP-1997: US-057777.
(HUMA-) HUMAN GENOME SCI INC.
Brewer LA, Carter KC, Dillion PJ, Ebner R, Endress GA,
Fan P, Feng P, Ferrile AM, Fischer CL, Florence C,
Florence K, Greene JM, Hu J, Kiyak H, Laflair DW, SM,
Li Y, Moore PA, N J, Olsen HS, Rosen CA, Rubin Z,
Shi Y, Soppet DR, Wei Y, Young P, Yu G, Zeng Z,
WPI: 99-059865/05.
DR N-PSDB: V84445.
PT New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
PS Claim 11: Page 506-507: 772pp; English.
CC The invention relates to nucleic acid sequences (V84411 to V84633)
CC encoding human secreted proteins (W8534 to W85756). The secreted protein
CC gene sequences are deposited with the ATCC under deposit numbers ATCC
CC 57979, 97974, 97975, 97976, 97977, 209007, 209008, 209009, 209010,
CC 209011, 209080, 209081, 209082, 209083, 209084, 209085, 209511. Host
CC cells comprising recombinant vectors containing the nucleic acid
CC sequences are used for the recombinant production of the secreted
CC proteins. The polynucleotide and amino acid sequences are useful for are
CC useful for preventing, treating or ameliorating medical conditions e.g.
CC by protein or gene therapy. Pathological conditions can be also diagnosed
CC by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the new polynucleotides, based on
CC which tissues they are most highly expressed in, and include developing
CC products for the diagnosis or treatment of cancer, neurodegenerative
CC disorders, developmental abnormalities and foetal deficiencies, blood
CC disorders, tumours, leukemias, diseases of the immune system, autoimmune
CC diseases, hepatic and renal diseases, lymphomas, inflammation, allergies,
CC ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,
CC restenosis, prostate diseases, obesity, disorders involving osteoclasts
CC or lymph, digestive/endocrine disorders, infections and AIDS. The
CC polypeptides are also useful for identifying their binding partners.
CC The present sequence represents human secreted protein (see descriptor
CC line for gene number and clone identification).
SQ Sequence 221 AA;

568	Length: 221	February 11, 2000 15:49	Type: P	Check: 9823
1	MHGHGATATGC	IGAPONCKEL	YEEKYLKLNNA	REBERXYDNNA
				ELFAVYKMQ
51	ALEKATATKDC	VSPSEYTYAC	SRLLVYKKA	FQVQSGEIS
				SIDECRRFR
101	LDCPLAMERI	KEDRITITKD	DKGLNLCIA	DVYSLEIYIM
				DKLREIYAM
151	DEIQIPIREL	METHMRSHL	PPDEGRQIV	SQMLQTLISM
				SASDELDDSQ
201	VROMFLDES	AYNAFNREL	A	

ID 11AA SEQUENCE 1.0
 AC W83391: Protein: 161 AA.
 DT 29-MAR-1999 (first entry)
 DE Caenorhabditis elegans synmv protein LIN-52.
 KW LIN-52; synthetic multivulvar; Synmv; signal transduction;
 KW animal model; tumour suppressor; retinoblastoma; cancer; cancer;
 KW cell proliferation; gene therapy.
 OS Caenorhabditis elegans.
 PN M09854299-A1.
 PD 03-DEC-1998.
 PF 28-MAY-1998: U11043.
 PR 28-MAY-1997: US-047966.
 RA (MAST) MASSACHUSETTS INST TECHNOLOGY.
 PI Ceol C, Horvitz HR, Lu X;
 DR WPI: 99-045362/04.
 DR N-PSDB: V72864.
 PT Novel LIN-37, -35, -55, -52, -53 and -54 gene from C. elegans -
 useful for treating diseases associated with altered levels of cell
 proliferation, e.g. carcinomas
 PR Claim 7: Fig 11: 70pp: English.
 CS This is the amino acid sequence of LIN-52, a novel protein of
 CC Caenorhabditis elegans. The lin-52 gene (see V72864) is a novel
 CC synthetic multivulvar (synmv) gene involved in cell fate and cell
 CC proliferation, and is part of a pathway for tumour suppression and
 CC genetic and biochemical model system for tumour suppression and
 CC cancer in mammals. Synmv pathway genes and proteins may be used
 CC to identify genes which are part of the mammalian pathway and
 CC to identify genes, proteins and therapeutic compounds which
 CC modulate this pathway. Pure nucleic acids (see V72859-65) encoding
 CC C. elegans synmv polypeptides selected from LIN-37, -35, -55, -52,
 CC -53, -54 and E2F-1 (see W83396-92), are new. Also claimed are:
 CC (1) nucleic acids (see V72865-67) encoding the mouse and human
 CC homologues of C. elegans LIN-54; (2) vectors containing the nucleic
 CC acids; (3) transgenic cells; (4) a pure mammalian synmv
 CC polypeptide, and (5) an antibody which binds to a synmv family
 CC protein. The synmv nucleic acids and polypeptides can be used to
 CC diagnose and treat, especially by gene therapy, conditions
 CC involving altered levels of cell proliferation, e.g. synmv-
 CC associated carcinomas.
 SQ Sequence 161 AA.

```

W83391 Length: 161 February 11, 2000 15:49 Type: P Check: 9212 ..

      1  MSRPGLFTGY EFGDDDEMFVQ OMTEKSNAE QAKMLEQOK MLECIETMPE
      51  ESEPYPMKCL DFEAFQSES VSKQYESPY NISFLKEDAV TWNTSHCPA
     101  DDIATLIRNI QNSVYTLGIE EAFQCRGRKL LNVLKPTGSA SPRVLOPTTP
     151  KNVAETGCS Q

!1AA SEQUENCE 1.0
ID   W95506 standard; peptide; 446 AA.
AC   W95506;
DT   26-MAR-1999 (first entry)
DE   Mortierella alpina delta 5 desaturase.
KW   Delta 5 desaturase; recombinant; fatty acid desaturase; FAD; PUFA; oil;
KW   polyunsaturated fatty acid; linoleic acid; arachidonic acid; linolenic;
KW   stearidonic acid; eicosapentaenoic acid; malnutrition; feeding formula;
KW   dietary supplement; prostaglandin; restenosis; angioplasty; inflammation;
KW   rheumatoid arthritis; psoriasis; osteoporosis; cancer; eczema; AIDS;
KW   diabetes; cosmetic; animal feed.
OS   Mortierella alpina.
SN   M09846764-01.
PD   22-OCT-1998.
PF   10-APR-1998; 007421.
PR   24-OCT-1997; US-956985.
PR   11-APR-1997; US-833610.
PR   11-APR-1997; US-834033.
PR   11-APR-1997; US-834655.
RA   (ABRO ) ABBOTT LAB.
PA   (CALT ) CALGENE LLC.

```

PI Chaudhary S, Huang Y, Knutson D, Leonard AE, Mukerji P,
PI Thummond J, WPI: 99-080739/07.
DR N-PSDB: X00891.
PT Nucleic acid construct able to express fatty acid desaturase in
PT plants - useful in human or animal nutrition, as cosmetics and
PT therapeutically, e.g. for testosterone, cancer and diabetes
PS Claim 7: Fig 7A-D: 210pp: English.
CC This represents a Mortierella alpina delta 5 desaturase. The invention
CC relates to a nucleic acid construct that contains at least one of the
CC nucleotide sequences (X00889 to X00891) encoding M. alpina delta 6,
CC delta 12 and delta 5 desaturases (W95504 to W95506) respectively, coupled
CC to an expression control sequence functional in plants. Recombinant plant
CC cells containing at least one DNA encoding a M. alpina fatty acid
CC desaturase (FAD), can be used for the production of polyunsaturated fatty
CC acids (PUFA). These recombinant cells of plants containing them are used
CC to produce oils such as linoleic acid, arachidonic acid, gamma-linolenic
CC acid, dihomogamma-linolenic acid, stearidonic acid and cis-eicosapentaenoic
CC acid (EPA). These plant oils are used: (i) to treat malnutrition, (ii) in
CC infant feeding formulas, or dietary supplements or substitutes, for use
CC in humans or animals, (iii) for treating disorders associated with such
CC inadequate consumption or production of PUFA (or their metabolites such
CC as prostaglandins), e.g. testosterone after angioplasty, cancer, eczema,
CC AIDS, rheumatoid arthritis, psoriasis, osteoporosis, inflammation,
CC diabetes; (iv) as cosmetics, and (v) as animal feeds. Fragments of the
CC DNA are used as probes to isolate related coding sequences. Recombinant
CC plants can produce high yields of PUFA, since new pathways can be created
CC and unwanted ones suppressed. e.g. plants can be engineered to express oils of
CC particular PUFA composition. e.g. one similar to that in human milk, and
CC product recovery is simpler than with e.g. fish.
CC Sequence 446 AA;

W95506 Length: 446 February 11, 2000 15:49 Type: P Check: 5663

1 MGNDQKFTI WEELAAHNTK DLLLAIAIGR YVDVTKFELR HPGGVKRIILL
51 GAGRDVTPVE EMYHAFGAAD AIKKRYVGT LVANEKPIPP EPTVEHNTIK
101 TRVEGETFDR NIDKRNPEI MGRVALIYGS LIASYSACLF VPFEVETWLL
151 QVAFALIMGF ACAQVGLNPL HDSASFVYTH NPATWKLTA THDEFNASTY
201 LVWNYGHMIG HHPTYNIGA DDDVSTSEPD VRRIKNOQM EVNHINOHMF
251 VPLVYGLIAF KVRIODINIL YEVKINDAIR VNPISHTWY MFWGKAFFV
301 WYRLIPILOY LPLGVKLLF TVADVSSYW LALTFQANW VEEVOWPLPD
351 ENGIIOKDMA AMQVETODY AHDSHLMTSI TGSINQAVH HLEPNVSOH
401 YPDILAIIN TCSEYKVPYL VKDTWQAFQ SHLEHLRVLG LRPKEE
11AA-SEQUENCE 1.0
ID W8882, standard; Protein; 304 AA.
DC W8882, 1999 (first entry)
DI 31-MAR-1999
DE H. pylori, GHPD 1670 protein.
KW GHPD protein, Helicobacter infection; gastroduodenal disease; gastritis;
KM peptic ulcer disease.
OS Helicobacter pylori.
PN W09843478-A1.
PD 08-OCT-1998.
PE 01-APR-1998; U06371.
PR 29-JUL-1997; US-502615.
PR 01-APR-1997; US-833457.
PR 24-JUN-1997; US-881227.
PA (HUMA-) HUMAN GENOME SCI INC.
PI (INMR) MERIEUX ORAVAX PASTEREUR MERIEUX SERUMS.
PI Al-Gazawi A, Kleantous H, Miller C, Oomen RP, Tomb J;
PI WPI: 98-542293/46.
DR N-PSDB; X14581.
DR New isolated Helicobacter polynucleotides - used to develop products
PT for the diagnosis, prevention and treatment of Helicobacter

PT Infections and gastrointestinal diseases.
 PS Claim 8, Page 1342-1943; 20549; English.
 CC This sequence represents a Helicobacter pylori GHPo protein of the
 CC invention. The polypeptides can be used for preventing or treating
 CC Helicobacter infections, and gastrointestinal diseases associated with
 CC these infections, including acute, chronic, and atrophic gastritis, and
 CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be
 CC used for the production of antibodies. The products can also be used for
 CC detection and diagnosis.
 CC Sequence 304 AA;
 CC

W98862 Length: 304 February 11, 2000 15:49 Type: P Check: 4

1 MESAALYCK EKAILDLVG VPKTYSMAOE CNILGLYDDA FIITROENLV

51 GIISLOGIST SNMOKDIEG YEPARONVUN TISKDIOURI VAKRRKEFIN

101 QSPNIDNIYA KAITQIESK GIKTEYELV PETITSNVS EPEKKKLEMT

151 TSINEELIES SKEDKOENEN XSNETHSNTS SKKDKKNNKF KKTTESTSK

201 RALLIQTIER YKNAIKEKP TLINSKEVTN FYAEINGKY IAPNPKLKL

251 SDTILHMCI LKRTLSLNF KIKTPEVRVA GLRLMRAPKF LNSLYLLEYT

301 PKLN

```

W98755 length: 102 February 11, 2000 15:49 Type: P Check: 3203
1 MISEKRTSP CWSVPVLSIK IFFETLARPSK NNSLNTVIG FLNPTTNAM
51 GVARANQOG LMTSLIHTI KAFSKLPKP NCNKKRLTKA KNTALEKIS
101 AI
!!AA=SEQUENCE 1.0
ID W98226 standard; Protein: 115 AA.
AC W98226;
DT 31-MAR-1999 (first entry)
H. pylori; GPO 1099 protein.
KW GPO protein; Helicobacter infection; gastroduodenal disease; gastritis;

```


KW peptic ulcer disease.
 OS Helicobacter pylori.
 PN W09843478-A1.
 PD 08-OCT-1998.
 PF 01-APR-1998; U06371.
 PR 29-JUL-1997; US-902615.
 PR 01-APR-1997; US-833457.
 PR 24-JUN-1997; US-881227.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
 DR Al-Garawi A, Kleantous H, Miller C, Oomen RP, Tomb J;
 WPI: 98-542293/46.
 DR N-PSDB: X13945.
 PT New isolated Helicobacter polynucleotides - used to develop products
 for the diagnosis, prevention and treatment of Helicobacter
 PT infections and gastrointestinal diseases
 PS Claim 8; Page 159-160; 2054pp; English.
 CC This sequence represents a Helicobacter pylori GMPD protein of the
 CC invention. The polypeptides can be used for preventing or treating
 CC Helicobacter infections, and gastroduodenal diseases associated with
 CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be
 CC used for the production of antibodies. The products can also be used for
 CC detection and diagnosis.
 OS Sequence 115 AA;

W98226 Length: 115 February 11, 2000 15:49 Type: P Check: 2117 ..

1 MSKTTHRIA ISFVLSLFA CDEPKSSQ SHQNTKTK NNPINQAND
 51 IRKIEHEED EKATKEVNDL INNENKIDEI NNEENADPSQ KTNINVLQRA
 101 TNDQNLNSP LNRKY

11AA_SEQUENCE 1.0
 ID W88330 standard; Protein; 336 AA.
 AC W88330.
 DC 06-APR-1999 (first entry)
 DE Mannosyl transferase involved in O antigen biosynthesis.
 KM O antigen; mannosyl transferase; wbm gene; diagnosis.
 OS Salimella enterica.
 PN W09850531-A1.
 PD 12-NOV-1998.
 PF 01-MAY-1998; AU0315.
 PR 22-JUL-1997; AU-008162.
 PR 01-MAY-1997; AU-005545.
 PA (UNSA-) UNIV SYDNEY.
 PI Reeves PJ, Wang L;
 DR W09850531/05.
 DR N-PSDB: X06750.
 PT Nucleic acid molecules specific for bacterial polysaccharide
 PT antigens or useful for detecting specific strains in, e.g. food,
 PT faeces or patient samples.
 PS This sequence, fig 9; 105pp; English.
 CC encoded by the amino acid sequence of a mannosyl transferase that is
 CC involved in the biosynthesis of a gene cluster (see X06750) involved in
 CC the biosynthesis of the Salmonella enterica serotype C2 strain M67
 CC O antigen. The use of nucleic acid molecules derived from
 CC assembly and transport genes, particularly wbd (transferase), wzx
 CC (flippase) and wzy (polymerase) genes, within O antigen gene
 CC clusters, improve the specificity of methods for the detection and
 CC identification of O antigens, e.g. in testing food or faecal
 CC derived samples, for samples from patients. The O antigen is a
 CC major virulence factor of S. enterica.
 OS Sequence 336 AA;

W88330 Length: 336 February 11, 2000 15:49 Type: P Check: 7734 ..

1 MNKKVLMDSI SWSNGGIGR FIDELSKLLC DISKELYRK CASPLAPGL
 51 AVNIFLKKTI DVFELPGIIP PLFCCKRTII TIHNLNIDL NDNSSLFKFL
 101 FYNIIIRGRC RKAYKIFIVS NFSKRIYAV SGVAPNKIIV VINGVSLFN

151 ADVKPINLIG KYLCVGNRK THNKECVIS AFAKADIDPS IKLVFTGNPC
 201 NDLEKLIICH GLSERKFFEG FVSEKDLPSL YKSGIGLVFP SLYGSGFLPV
 251 VEGMACGIPIY LITLSLSIPE VAGDAALIVD PLEDAITG ISRLINSEL
 301 RKHLIOGILL RAKRFNMQNV VSEIEWVLTE ACDGKN

11AA_SEQUENCE 1.0
 ID Y01440 standard; Protein; 64 AA.
 AC Y01440.
 DC 18-MAY-1999 (first entry)
 DE Secreted protein encoded by gene 58 clone HOD425.
 KM Human; secreted protein; gene therapy; protein therapy; cancer; weight;
 KM tumour; chromosome mapping; forensic; hematological disease; allergy;
 KM inflammation; cell proliferation; viral infection; wound healing;
 KM modulation; appetite; behaviour; food additive; preservative.
 OS Homo sapiens.
 PN W09303990-A1.
 PD 28-JAN-1999.
 PF 15-JUL-1998; U14613.
 PR 18-AUG-1997; US-056361.
 PR 16-JUL-1997; US-052661.
 PR 16-JUL-1997; US-052870.
 PR 16-JUL-1997; US-052871.
 PR 16-JUL-1997; US-052872.
 PR 16-JUL-1997; US-052873.
 PR 16-JUL-1997; US-052874.
 PR 16-JUL-1997; US-052875.
 PR 22-JUL-1997; US-053440.
 PR 22-JUL-1997; US-053441.
 PR 22-JUL-1997; US-053442.
 PR 18-AUG-1997; US-055683.
 PR 18-AUG-1997; US-055724.
 PR 18-AUG-1997; US-055725.
 PR 18-AUG-1997; US-055726.
 PR 18-AUG-1997; US-055946.
 PR 18-AUG-1997; US-055952.
 PR 18-AUG-1997; US-055985.
 PR 18-AUG-1997; US-055989.
 PR 18-AUG-1997; US-056359.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Duan R, Feng P, Ferlie AM, Florence KA, Fouad J, PE,
 PI Greene JM, Hu J, Ni J, Rosen CA, Ruben SM, Young JE,
 PI Yu G;
 DR WPI: 99-132234/11.
 DR N-PSDB: X22268.
 PT New nucleic acids encoding secreted human proteins - potentially
 PT useful for treating and diagnosing diseases and identifying specific
 PT binding agents.
 PS Claim 11; Page 232; 251pp; English.
 CC The invention relates to nucleic acid sequences (X22211 to X22282)
 CC encoding human secreted proteins (Y01383 to Y01454). The secreted protein
 CC gene sequences are deposited with the ATCC under deposit number ATCC
 CC 209138, 209139 or 209141. Host cells containing vectors comprising the
 CC nucleic acid sequences are used for the recombinant expression of the
 CC secreted proteins. The polynucleotide and amino acid sequences are useful
 CC for preventing, treating or ameliorating medical conditions e.g. by
 CC protein or gene therapy. Pathological conditions can be also diagnosed by
 CC determining the amount of the new polypeptides in a sample or by the
 CC presence of mutations in the new polynucleotides. The nucleic acid
 CC sequences, or its fragments, are useful for chromosome identification and
 CC mapping; as antisense and triplex-forming therapeutics; in gene therapy;
 CC for (forensic) identification of individuals; as molecular weight
 CC markers; to identify related sequences or specific mRNAs; in preparation
 CC of oligomers and to raise anti-DNA antibodies. Antibodies are useful as
 CC immunosassay reagents (including for in vivo imaging) and therapeutically
 CC to inhibit or activate particular polypeptides. A very wide range of
 CC disorders may be treated with the polynucleotide and polypeptide
 CC sequences, e.g. autoimmune or haematological diseases, allergy,
 CC inflammation, cancer or other forms of cell proliferation, viral or
 CC other infections. The sequences may also be useful in wound healing, to

CC modulate differentiation of embryonic stem cells, to modulate weight,
CC appetite, behaviour etc. and as food additive or preservative. The
CC present sequence represents a human secreted protein (see descriptor
CC line for gene number and clone identification).
CC Sequence 64 AA;
SQ

Y01440 Length: 64 February 11, 2000 15:49 Type: P Check: 9731 ..

1 MAASVGRATR SAAAHITQLP PAPPARTSP AOPDEGRKD ADWFRGPTV

51 NRTGSIQRL RGWA

11AA SEQUENCE 1.0
ID M99185 standard; Protein: 314 AA.

AC M99185;

DT 19-MAY-1999 (first entry)

DE Rhodococcus corallina cnp operon; biosensor; mycolic acid bacteria;

KW Rhodococcus corallina; cnp operon; biosensor; mycolic acid bacteria;

KV inducible promoter; environmental pollutant; industry; medicine.

OS Rhodococcus corallina.

PN MO9900517-A2.

PD 07-JAN-1999.

PF 29-JUN-1998; G01893.

PR 27-JUN-1997; GB-013666.

PT (UYCA) UNIV CAMBRIDGE TECH SERVICES LTD.

PI (UYCA) JAC, POWELL JAC, Roland HJ, Summers DK;

DR WPI139-055760/08.

DR N-PSDB; X19362

PT Isolating DNA encoding inducible promoter from mycolic acid bacteria

PT - useful to produce mycolic acid bacterial biosensors for particular

PT analytes, such as environmental pollutants, e.g. from industry or

PT medicine.

PS Example 7; Fig 4: 67pp; English.

CC A method has been developed for identifying and/or isolating DNA from

CC mycolic acid bacteria which encodes an inducible promoter induced in

CC response to a particular analyte (and/or associated operon proteins).

CC The method comprises: (a) culturing source of mycolic acid bacteria in

CC selective media containing specific analyte and selective for

CC oligotrophic bacteria; (b) identifying mycolic acid bacteria subsisting

CC on medium; (c) extracting DNA from these bacteria; (d) incorporating DNA

CC into vector; (e) cloning vector into suitable host cell; and (f)

CC screening host cells for inducible promoter and/or proteins to identify

CC vectors encoding it. The method allows isolation of DNA encoding a

CC promoter which is induced in response to a particular analyte (and/or

CC associated operon proteins), which can be used to produce biosensors for

CC the analyte. The DNA is incorporated into vectors which are used to

CC transform host cells to produce the biosensor. It is especially useful

CC when the analyte is an environmental pollutant (e.g. from industry or

CC medicine), especially a hydrophobic organic compound such as components of

CC fuels, pesticides; the biosensors may then be used to detect small

CC concentrations of analytes in samples e.g. for pollution monitoring. The

CC identified DNA can also be modified to produce modified inducible

CC promoters and/or operons also useful in biosensor production. The method

CC allows rapid isolation of promoters (and/or operon proteins) whilst

CC minimising host restrictions (e.g. thick cell walls of mycolic acid

CC bacteria which confer resistance to cell lysis) and requiring no

CC knowledge of inducible enzyme chemistry involved (which may be unknown

CC for specific analyte). The present sequence represents a protein from

CC the R. corallina cnp operon, given in the present invention.

CC Sequence 314 AA;

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

251 QIDAWNDWF VEOAGHSHE VRTWIAAYPA MSAKRYNT STEYREIHEM

301 IAGGITTAIV AVDE

11AA SEQUENCE 1.0
ID W97722 standard; Protein: 302 AA.

AC W97722;

DT 21-MAY-1999 (first entry)

DE Staphylococcus aureus mutant p15B9 virulence gene product.

KW Virulence p15B9; vaccine; antibacterial; antibiotic; screening;

KV bacteremia.

OS Staphylococcus aureus.

PN WO9901473-A2.

PD 14-JAN-1998; G01974.

PF 03-JUL-1998; US-867534.

PR (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

PI Holden DW;

DR WPI; 99-105999/09.

DR N-PSDB; X07134.

PT Inhibition of virulence genes from Staphylococcus aureus - useful

PT for, e.g. screening for potential anti-microbial agents

PS Claim 1; Page 179-180; 20pp; English

CC This is the deduced amino acid sequence of the protein product of a

CC mutant p15B9. S. aureus mutants were generated, and those with

CC of a signature-tagged transposon were identified in a mouse model of bacteraemia.

CC The nucleotide sequences of regions flanking the transposon insertion

CC sites of these mutants were determined, and database sequence

CC comparisons were performed to identify these virulence genes (see

CC X07088-136) and the possible function of their protein products

CC (see W97680-724). The p15B9 virulence gene product is unknown but

CC shows 41% identity to the hypothetical 33.7 kDa protein YJCT of

CC Bacillus subtilis p54604. A claimed method of identifying an

CC antibacterial agent involves assaying potential agents for the

CC ability to interfere with the expression of S. aureus virulence

CC gene products. Also new is a S. aureus organism containing a

CC functional mutation in one of the virulence genes, and its use in

CC vaccine compositions.

CC Sequence 302 AA;

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

W97722 Length: 302 February 11, 2000 15:49 Type: P Check: 1929 ..

1 MNOKRGLEI LINOKDQGM LFSVLKTKALK ASKRYIOWM SHQIKVNE

51 SYLNWYAK GDRYFIDLOE SEASSVPEY GERDILFEDN HMLINKPAG

101 IATHPNEDQO TGTIANIAI HYQINGETCK VRAVRLDOD TSGAIVAKH

151 RLAHAILDOQ LEKTKIRTY TAAEGKURT KKGINPPIG RDRSHPTRR

201 VSPGGTAVT HFYMASNAK ERLSIVLEL ETGTHQIRY HIASIGHPLT

251 GDSLYGGSK LINRQALHAN KQONVHPITD ELIYAEAPF ADMKNICRTY

301 FS

301 FS

11AA SEQUENCE 1.0
ID Y11009 standard; Protein: 142 AA.

AC Y11009;

DT 08-JUN-1999 (first entry)

DE H. pylori ORF hp622217_23564012_61.5 secreted protein.

KW Vaccine; probe; diagnostic; ORF; cell envelope protein;

KV secreted protein; cellular protein.

OS Helicobacter pylori.

PN WO9818323-A1.

PD 07-MAY-1998.

PF 28-OCT-1997; U19575.

PR 14-JUL-1997; US-891928.

PR 28-OCT-1996; US-739150.

PR 06-DEC-1996; US-759739.

PT (ASTR) ASTRA AB.

PT

PT

PT

PT

PT

PT

PT

PI Alm RA, Smith D;
DR WPI: 98-271811/24.
DR N-PSDB: X30476.
PT Helicobacter pylori nucleic acids and proteins - used to develop
PT products for the detection, prevention and treatment of H. pylori
infections
PS Claims 27, 31: Page 219: 27pp: English.
CC Recombinant or substantially pure preparations of H. pylori polypeptides
are disclosed, together with the nucleic acids encoding them. In all,
CC 73 ORFs are shown. The proteins are variously cell envelope proteins,
CC secreted proteins or other cellular proteins. Vaccines containing the
CC nucleic acids or proteins are claimed, as are probes containing at least
CC 8 nucleotides from the nucleic acid sequences. The vaccines are useful
CC for treating or reducing the risk of H. pylori infections, and the
CC probes can be used diagnostically for detecting the presence of
CC Helicobacter in a sample. The products are also of use in screening
CC for compounds having the ability to interfere with the H. pylori life
CC cycle or to inhibit H. pylori infection.
SQ Sequence 142 AA;
Y11009 Length: 142 February 11, 2000 15:49 Type: P Check: 2840 ..
1 MKTNFKIKL LFAMCLITGM FNAFLNADON TDIDISPD MALNSGLVS
51 RDOLKIEIRK ETLERKVAIL NDYNDKNVNI KEDDISLGSF QPNDNIGINA
101 MGIQNLIMS QMNSNYGPNN SFMYGAPTY SDSSELPIL GY
11AA_SEQUENCE 1.0
ID Y11062 standard; Protein: 116 AA.
AC Y11062:
DT 08-JUN-1999 (first entry)
DE H. pylori ORF 06cp30603_10744075_C3_136 secreted protein.
KW Vaccine; probe; diagnostic; ORF; cell envelope protein.
KM secreted protein; cytoplasmic protein; cellular protein.
NS Helicobacter pylori.
PN W09824475-A1.
PD 11-JUN-1998.
PF 05-DEC-1997: U22104.
PR 14-JUL-1997: US-891928.
PR 05-DEC-1996: US-759625.
PR 25-MAR-1997: US-823745.
PA (ASTRA) ASTRA AB
PI Alm RA, Castriotta LM, Doig PC, Kabok Z, Smith D;
DR WPI: 98-333051/29.
DR N-PSDB: X30591.
PT New isolated Helicobacter pylori nucleic acids - used to develop
PT products for the diagnosis, prevention and treatment of infection by
PT H. pylori and other Helicobacter species
PS Claims 37, 41: Page 224-225: 33pp: English.
CC Recombinant or substantially pure preparations of H. pylori polypeptides
are disclosed, together with the nucleic acids encoding them. In all,
CC 97 ORFs are shown. The proteins are variously cell envelope proteins,
CC cytoplasmic proteins, secreted proteins or other cellular proteins.
CC Vaccines containing the nucleic acids or proteins are claimed, as are
CC probes containing at least 8 nucleotides from the nucleic acid
CC sequences. The vaccines are useful for treating or reducing the risk of
CC H. pylori infections, and the probes can be used diagnostically for
CC detecting the presence of Helicobacter in a sample. The products are
CC also of use in screening for compounds having the ability to interfere
CC with the H. pylori life cycle or to inhibit H. pylori infection.
SQ Sequence 116 AA;
Y11062 Length: 116 February 11, 2000 15:49 Type: P Check: 4555 ..
1 LMSKITHFI VISFLVSLVS ACDEPKKSS OSHQNTKTT QNNQINPNK
51 DIKIEHEEE DEKTKREVND LINNEKIDE INNERNAPS OKRTNNVLOR
101 ATNHQDNLSS PLNRRY
11AA_SEQUENCE 1.0
ID W97887 standard; Protein: 375 AA.

AC W97887;
DT 07-JUN-1999 (first entry)
DE Human myostatin.
KW Myostatin; human; transforming growth factor beta;
KW double muscling; muscle hyperplasia; transgenic animal.
OS Homo sapiens.
PN W09902667-A1.
PD 21-JAN-1999.
PR 14-JUL-1998: IB1197.
PR 15-JAN-1998: US-007761.
PR 14-JUL-1997: US-891789.
PA (UyLi-) UNIV LIEGE.
PI Georges M, Grobet L, Poncelet D;
DR WPI: 99-120869/10.
PT Increasing muscle mass in mammals - by decreasing myostatin
PT expression
PS Disclosure; Page 65: 75pp: English.
CC This is the amino acid sequence of human myostatin, a member of
CC the transforming growth factor-beta superfamily. The invention
CC relates to factors affecting muscle development in mammals,
CC especially to cloning of the myostatin gene and determining
CC the role of the gene in muscle hyperplasia. A mutation of the
CC bovine myostatin gene (see X24415-16) has been detected. Certain
CC breeds of cattle homozygous for the mutant gene are double-muscled.
CC A new method of increasing muscle mass of a mammal having myostatin-
CC expressing muscle cells, comprises administration of a nucleic acid
CC molecule substantially complementary to at least a portion of mRNA
CC encoding myostatin and of sufficient length to reduce myostatin
CC expression and thus increase muscle mass. A ribozyme may also be
CC used. Also claimed are: a method for determining muscular
CC hyperplasia (MH) in a mammal using primers based upstream and
CC downstream of the mutation; a diagnostic kit for determining the
CC genotype of a sample of genetic material; a method for determining
CC MH in a mammal; a method for determining double muscling in a
CC bovine; a method for determining the myostatin genotype of an
CC animal; purified myostatin; isolated nucleic acids; a microbial
CC host cell; a probe based on the myostatin gene mutation; transgenic
CC mammals having MH phenotype; and a myostatin knockout animal.
CC Primers are preferably based on genomic bovine myostatin DNA (see
CC X24464) and human myostatin cDNA (see X24418).
SQ Sequence 375 AA;
W97887 Length: 375 February 11, 2000 15:49 Type: P Check: 1814 ..
1 MOKLQCVYI YLFMLIVAGP VDLNENSEOK ENVEKGLCN ACTWRONTKS
51 SRIPAKIQI LSKRLIETAP NISKDVIRQL LPRAPLREL IDQYVORPD
101 SSDSLEDDE YHATTEIIT MPTSEFLMQ VDGKPKCFE KESSKIQYNK
151 VVKQALWYL RPYEPPTVF VOILRLIKPM KQSTRYTGIR SLKIDMNPST
201 GIMQSDIKT VLQNLKQPE SNLGEIKAL DENGHDLAFT FPGGEDGLN
251 PFEVKTDT PKRSRDEFL DCDHSTESH CCRYPITVDF EAFQMDWITA
301 PKRYKATCS GECEYFLOK YPHHLVHQQ NRRGSGPCC TPTMSPLIM
351 LYNGKEQIY YGKIPAVVD RCGCS
11AA_SEQUENCE 1.0
ID W97884 standard; Protein: 375 AA.
AC W97884;
DT 07-JUN-1999 (first entry)
DE Bovine myostatin.
KW Myostatin; cattle; bovine; transforming growth factor beta;
KW double muscling; muscle hyperplasia; transgenic animal.
OS Bos taurus.
PN W09902667-A1.
PD 21-JAN-1999.
PR 14-JUL-1998: IB1197.
PR 15-JAN-1998: US-007761.
PR 14-JUL-1997: US-891789.

PA (UYLI-) UNIV LIEGE.
 PI Georges M, Grobet L, Poncetlet D;
 DR WPI: 99-120869/10.
 DR N-PSDB: X24415, X24464.
 PT Increasing muscle mass in mammals - by decreasing myostatin
 expression
 PS Claim 19: Page 55; 75pp; English.
 CC This is the amino acid sequence of bovine myostatin, a member of
 CC the transforming growth factor beta superfamily, as encoded by the
 CC wild-type gene (see X24415). A mutation of this gene (see X24416)
 CC has been detected in cattle. Cattle of the Belgian Blue breed
 CC homozygous for the mutant gene are double-muscling. A new method of
 CC increasing muscle mass of a mammal having myostatin-expressing
 CC muscle cells, comprises administration of a nucleic acid molecule
 CC substantially complementary to at least a portion of mRNA
 CC encoding myostatin and of sufficient length to reduce myostatin
 CC expression and thus increase muscle mass. A ribozyme may also be
 CC used. Also claimed are: a method for determining muscular
 CC hyperplasia (MH) in a mammal using primers based upstream and
 CC downstream of the mutation; a diagnostic kit for determining
 CC the genotype of a sample of genetic material; a method for
 CC determining MH in a mammal; a method for determining double
 CC muscling in a bovine animal; a method for determining the myostatin
 CC genotype of an animal; purified myostatin; isolated nucleic acids;
 CC a microbial host cell; a probe based on the myostatin gene
 CC mutation; transgenic mammals having MH phenotype; a myostatin
 CC knockout animal; and a heterologous nucleotide sequence antisense
 CC to that gene, and optionally further containing a gene encoding a
 CC nucleic acid sequence with ribozyme activity in transcriptional
 CC association with the antisense sequence.
 SQ Sequence 375 AA;
 W97884 Length: 375 February 11, 2000 15:49 Type: P Check: 8371 ..

1 MOKLQISYVI YLFMLIVAGP VDLNENSEOK ENVEKEGICN ACIMRENTTS
 51 SRELEAIKIOI LSKULEETAP NISKDAIROL LPKAPLLEL IDQFDVORDA
 101 SSDGSELEDD YHARETVIT MPTESDLITQ VEGKPKCCFF KFSKSIQYNK
 151 LKRAQLMIYL RPKVTPATVF VOILRLIKPM KDGRYTGIR SKLDMPNPT
 201 GIGOSIDVKT VLQWMLKQPE SNLGEIKAL DENGHDIAVT FEPGEGDILT
 251 PLEVKVYDT PKRSRDEGL DCDERSTESR CCRPLTVDF EAFGMWIIA
 301 PRRYANCOS GECEVFLOK YPHILVHOA NPKASAPCC TTKMSPIM
 351 LYFNGEGOII YGKIPAMVVD RCGCS

11AA_SEQUENCE 1.0
 ID W97885 standard; Protein; 286 AA.
 AC W97885;
 DT 07-JUN-1999 (first entry)
 DE Bovine myostatin (mutant form).
 KW Myostatin; cattle; bovine; transforming growth factor beta;
 KM double muscling; muscle hyperplasia; transgenic animal; mutant.
 OS Bos taurus.
 PN WO9902667-A1.
 PD 21-JAN-1999.
 PE 14-JUL-1998; IB1197.
 PF 15-JAN-1998; US-007761.
 PR 14-JUL-1997; US-891789.
 PA (UYLI-) UNIV LIEGE.
 PI Georges M, Grobet L, Poncetlet D;
 DR WPI: 99-120869/10.
 DR N-PSDB: X24416.
 PT Increasing muscle mass in mammals - by decreasing myostatin
 PT expression
 PS Disclosure; Page 57; 75pp; English.
 CC This is the amino acid sequence of a mutant bovine myostatin,
 CC which is truncated when compared to the wild-type polypeptide (see

CC W97884). The mutation results from an 11-bp deletion, i.e.
 CC nt82del(11), in the coding region (see X24416) of the myostatin
 CC gene. Cattle of the Belgian Blue breed homozygous for the mutant
 CC gene are double-muscling. A new method of increasing muscle mass of
 CC a mammal having myostatin-expressing muscle cells, comprises
 CC administration of a nucleic acid molecule substantially
 CC complementary to at least a portion of mRNA encoding myostatin and
 CC of sufficient length to reduce myostatin expression and thus
 CC increase muscle mass. Also claimed are: a method for determining
 CC muscular hyperplasia (MH) in a mammal using primers based upstream
 CC and downstream of the mutation; a diagnostic kit for determining
 CC the genotype of a sample of genetic material; a method for
 CC determining MH in a mammal; a method for determining double
 CC muscling in a bovine animal; a method for determining the myostatin
 CC genotype of an animal; purified myostatin; isolated nucleic acids;
 CC a microbial host cell; a probe based on the myostatin gene
 CC mutation; transgenic mammals having MH phenotype; a myostatin
 CC knockout animal; and a heterologous nucleotide sequence antisense
 CC to that gene, and optionally further containing a gene encoding a
 CC nucleic acid sequence with ribozyme activity in transcriptional
 CC association with the antisense sequence.
 SQ Sequence 286 AA;
 W97885 Length: 286 February 11, 2000 15:49 Type: P Check: 1462 ..

1 MOKLQISYVI YLFMLIVAGP VDLNENSEOK ENVEKEGICN ACIMRENTTS
 51 SRELEAIKIOI LSKULEETAP NISKDAIROL LPKAPLLEL IDQFDVORDA
 101 SSDGSELEDD YHARETVIT MPTESDLITQ VEGKPKCCFF KFSKSIQYNK
 151 LKRAQLMIYL RPKVTPATVF VOILRLIKPM KDGRYTGIR SKLDMPNPT
 201 GIGOSIDVKT VLQWMLKQPE SNLGEIKAL DENGHDIAVT FEPGEGDILT
 251 PLEVKVYDT PKRSRDEGL DCDERSTESR CCRPLTVDF EAFGMWIIA
 301 PRRYANCOS GECEVFLOK YPHILVHOA NPKASAPCC TTKMSPIM
 351 LYFNGEGOII YGKIPAMVVD RCGCS

11AA_SEQUENCE 1.0
 ID W97886 standard; Protein; 376 AA.
 AC W97886;
 DT 07-JUN-1999 (first entry)
 DE Murine myostatin.
 KW Myostatin; mouse; transforming growth factor beta;
 KM double muscling; muscle hyperplasia; transgenic animal.
 OS Mus sp.
 PN WO9902667-A1.
 PD 21-JAN-1999.
 PE 14-JUL-1998; IB1197.
 PF 15-JAN-1998; US-007761.
 PR 14-JUL-1997; US-891789.
 PA (UYLI-) UNIV LIEGE.
 PI Georges M, Grobet L, Poncetlet D;
 DR WPI: 99-120869/10.
 DR N-PSDB: X24417.
 PT Increasing muscle mass in mammals - by decreasing myostatin
 PT expression
 PS Disclosure; Page 60; 75pp; English.
 CC This is the amino acid sequence of murine myostatin, a member of
 CC the transforming growth factor beta superfamily. The invention
 CC relates to factors affecting muscle development in mammals,
 CC including the detection of a mutation in the bovine myostatin
 CC gene (see X24415-16). Cattle of the Belgian Blue breed homozygous
 CC for the mutant gene are double-muscling. A new method of increasing
 CC muscle mass of a mammal having myostatin-expressing muscle cells,
 CC comprises administration of a nucleic acid molecule substantially
 CC complementary to at least a portion of mRNA encoding myostatin
 CC (including murine myostatin) and of sufficient length to reduce
 CC myostatin expression and thus increase muscle mass. A ribozyme may
 CC also be used. Also claimed are: a method for determining muscular
 CC hyperplasia (MH) in a mammal using primers based upstream and
 CC downstream of the mutation; a diagnostic kit for determining
 CC the genotype of a sample of genetic material; a method for

CC determining MH in a mammal; a method for determining double
CC muscling in a bovine animal; a method for determining the myostatin
CC genotype of an animal; purified myostatin; isolated nucleic acids;
CC a microbial host cell; a probe based on the myostatin gene
CC mutation; transgenic mammals having MH phenotype; and a myostatin
CC knockout animal; and a transgenic bovine having a gene encoding
CC active myostatin.
SO Sequence 376 AA;

W9786 Length: 376 February 11, 2000 15:49 Type: P Check: 2420

1 MMOKLMYV IYLFMLIAG PVDLNEGSR EENYERKGLC MCAMWQNR
51 YSRFAIKIQ ILSKRLETA PNISKDAIRQ LLPRAPPLRE LIDQYDVHRD
101 DSSDGSLEDD DYHATETETII TMPRESDFLM QADGKPKCF FFFSKIQYN
151 KVKQAQIMY LRPVKTPTTV FVQILRLIKP MKDGTATGI RSLKIDMSPG
201 TGIWQSIDVR TVLQNLKOP ESNLGEIEKA LDENGHDLAV TTPGPEDEL
251 NPFLEAVTD TPKRSRDFG LDCDEHSTES RCGRPLTVD FEAFGDWII
301 APRKYKANYC SGCEFEVFIQ KYPHHLVHQ ANPRGSAGPC CPTKMSPIN
351 MLVFNKKEOI IYKIPAMV DRSGCS
11AA_SEQUENCE 1.0
ID W97840 standard; Protein: 207 AA.
AC W97840: 1999 (first entry)
DE Winged bean chymotrypsin inhibitor WCI-3.
KW Chymotrypsin inhibitor; WCI-3; trypsin inhibitor; WTI-1B;
KM protease inhibitor; enzyme engineering; protein engineering;
KW artificially synthesised gene; winged bean; rice; transgenic plant;
OS insect resistance.
PS Psophocarpus tetragonolobus.
FH Key Location/Qualifiers
FT 1..24
FT Peptide 1..24
FT /note= "hydrophilic peptide utilised in
FT synthetic trypsin inhibitor"
PD EP-900842-A2.
PN 10-MAR-1999.
PR 01-SEP-1998: 307004.
PR 01-SEP-1997: JP-236332.
PA (NORO) TOHOKU NAT AGRIC EXPERIMENT STATION.
PI Mochizuki A;
DR MPI: 99-155935/14.
DR N-PSDB: X24283.
PT Artificial synthetic trypsin inhibitor gene (WTI-1b) useful for
PT insect resistance in transgenic plants - designed on winged bean
PT chymotrypsin inhibitor (WCI-3) using codon and base comparisons
PS Disclosure: Page 15: 21pp: English.
CC This polypeptide comprises chymotrypsin inhibitor WCI-3 of winged
CC bean (Psophocarpus tetragonolobus (L.) DC.). An artificially
CC synthesised gene (see X24282) is claimed which codes for a trypsin
CC inhibitor (see W97839) designed for stable expression in rice.
CC The trypsin inhibitor comprises the winged bean WTI-1b polypeptide
CC having an N-terminal hydrophilic peptide derived from WCI-3 (i.e.
CC amino acid residues 1-24 of WCI-3) utilised for transporting the
CC protein precursor to a vacuole after translation. The synthetic
CC was obtained by: (i) comparing WTI-1b and WCI-3 amino acid
CC sequences, and obtaining homologous amino acid pairs (one each
CC from WTI-1a and b); and (ii) selecting codons from WCI-3 if the
CC amino acid pair is the same, and from WTI-1b if they are
CC different. High usage codons are selected and sequences which
CC cause mRNA stability are modified (see also X24294-96). Also
CC claimed is a plant containing the artificially synthesised gene.
CC Transgenic plants can be produced that are resistant to pests,
CC especially lepidopteran insects.
SO Sequence 207 AA;

W97840 Length: 207 February 11, 2000 15:49 Type: P Check: 7697

1 MKSTFLALF LSAIISHLP SSTADDDLYD AEGNLVENG TYLLPHIWA
51 HGGGIETAKT GNEPCPLTV RSPNEVSKGE PIRISSQFLS LFIPRGSIVA
101 LGFANPSCA ASPMWTVDVDS PGPAVKLSQ QKLEKIDLV FFEKXSHSN
151 IHVYKLAYCO HDEEDVQCDQ YIGIHRDRNG NRRLVATEEN PLBLVLKAK
201 SETASSH

11AA_SEQUENCE 1.0
ID Y03794 standard; Protein: 144 AA.
AC Y03794:
DE 11-JUN-1999 (first entry)
KW S. aureus polypeptide.
KW Staphylococcus aureus polypeptide; thyroiditis; infective carditis;
KW lung abscess; secretory diarrhoea; cerebral abscess; conjunctivitis;
KW toxic shock syndrome; folliculitis; septic arthritis; antibacterial;
KW H. pylori infection; gastric ulcer; adenocarcinoma.
OS Staphylococcus aureus.
PN EP-905243-A2.
PD 31-MAR-1999.
PF 03-AUG-1998: 306185.
PR 05-AUG-1997: US-055387.
PA (SMIR) SMITHKLINE BEECHAM CORP.
PA (SMIR) SMITHKLINE BEECHAM PLC.
PI Burnham MKR, Lonetto MA, Warren PV;
DR MPI: 99-192667/17.
DR N-PSDB: X31864.
PT New essential polypeptides from Staphylococcus aureus useful for
PT treating diseases such as infective endocarditis and toxic shock
PT syndrome
PS Claim 31: Page 61-62: 70pp: English.
CC The invention provides new Staphylococcus aureus polypeptides (Y03781-94)
CC and the genes (X31851-864) encoding them. Host cells containing vectors
CC comprising the nucleic acid sequences are used for the recombinant
CC expression of the proteins. The polypeptides can be used to screen for
CC modulators for use in antibacterial therapy. The polypeptides, their
CC antagonists and agonists are used to prevent or treat diseases caused by
CC S. aureus such as thyroiditis, lung abscesses, infective carditis,
CC secretory diarrhoea, cerebral abscesses, conjunctivitis, toxic shock
CC syndrome, folliculitis and septic arthritis. Screening for the presence of
CC the polypeptides may be used to diagnose, predict the susceptibility to,
CC or stage the progress of these S. aureus diseases and diseases caused by
CC Helicobacter pylori such as gastric ulcers and gastric adenocarcinoma.
CC There is not much information known about the essential genes expressed
CC by S. aureus during infection but these new polypeptides have been
CC identified as essential. They can therefore be used to develop
CC antibacterial compounds specific for those essential genes and this
CC ensures the effectiveness of the compounds in killing S. aureus. In
CC addition, these polypeptides can be used to effectively diagnose and
CC treat infections and diseases caused by S. aureus without the risk of
CC development of antibiotic resistance. The present sequence represents a
CC S. aureus polypeptide which has homology to an apparent transcriptional
CC regulator.
SO Sequence 144 AA;

Y03794 Length: 144 February 11, 2000 15:49 Type: P Check: 7277

1 MDRTKSLNV FVGNNRALDT LEOITKEDVK RYGINITEFA VLBLEYKNGP
51 PQIRITRDV LIASSISYV VSOLEDKWV TREKDKDKR VYMACLTEKG
101 QSQMADIFPK HAEITKRAFV VTKDELTLT QQARKLSAQ STEV
11AA_SEQUENCE 1.0
ID Y03782 standard; Protein: 135 AA.
AC Y03782:
DE 11-JUN-1999 (first entry)
KW S. aureus polypeptide.
KW Staphylococcus aureus polypeptide; thyroiditis; infective carditis;
KW lung abscess; secretory diarrhoea; cerebral abscess; conjunctivitis;

KW toxic shock syndrome; folliculitis; septic arthritis; antibacterial;
 KW H pylori infection; gastric ulcer; adenocarcinoma.
 OS Staphylococcus aureus.
 PN EP-905243-A2.
 PD 31-MAR-1999.
 PR 03-AUG-1998; 306185.
 PR 05-AUG-1997; US-055387.
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PI Burnham MKR, Lonetto MA, Warren PV;
 DR WPI: 99-192667/17.
 N-PSDB: X31852.
 PT New essential polypeptides from Staphylococcus aureus useful for
 PT treating diseases such as infective endocarditis and toxic shock
 PT syndrome
 PS Claim 11: Page 24-25; 70pp; English.
 CC The invention provides new Staphylococcus aureus polypeptides (Y03781-94)
 CC and the genes (X31851-864) encoding them. Host cells containing vectors
 CC comprising the nucleic acid sequences are used for the recombinant
 CC expression of the proteins. The polypeptides can be used to screen for
 CC modulators for use in antibacterial therapy. The polypeptides, their
 CC antagonists and agonists are used to prevent or treat diseases caused by
 CC S. aureus such as chyloditis, lung abscesses, infective carditis,
 CC secretory diarrhoea, cerebral abscesses, conjunctivitis, toxic shock
 CC syndrome, folliculitis and septic arthritis. Screening for the presence of
 CC the polypeptides may be used to diagnose, predict the susceptibility to,
 CC or stage the progress of these S. aureus diseases and diseases caused by
 CC Helicobacter pylori such as gastric ulcers and gastric adenocarcinoma.
 CC There is not much information known about the essential genes expressed
 CC by S. aureus during infection but these new polypeptides have been
 CC identified as essential. They can therefore be used to develop
 CC antibacterial compounds specific for those essential genes and this
 CC ensures the effectiveness of the compounds in killing S. aureus. In
 CC addition, these polypeptides can be used to effectively diagnose and
 CC treat infections and diseases caused by S. aureus without the risk of
 CC development of antibiotic resistance. The present sequence represents a
 CC S. aureus polypeptide which has homology to a B. subtilis probable
 CC acetyltransferase.
 CC Sequence 135 AA;
 SQ
 Y03782 Length: 135 February 11, 2000 15:49 Type: P Check: 881 ..
 1 INHRAFIKRLGLDQADGNGFOCGIWEYEGT LVGIGLHEI NHHKRTSIG
 51 YLYKQFEGH GIMTQALEAL INVCYKPYF ISITITASFH FLINSTYFMT
 101 RDLDPHSIF IFPSHSTWL FISTTHLSL LIFSF
 IIA-SEQUENCE 1.0
 ID Y02934 standard; Protein: 53 AA.
 AC Y02934;
 DT 11-JUN-1999 (first entry)
 DE Fragment of human secreted protein encoded by gene 106.
 KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;
 KW immune system; asthma; lymphocytic disease; brain; hepatitis; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 OS Homo sapiens.
 PN WO9902546-A1.
 PD 21-JAN-1999.
 PD 07-JUL-1998; U13684.
 PR 12-SEP-1997; US-058785.
 PR 08-JUL-1997; US-051916.
 PR 08-JUL-1997; US-051918.
 PR 08-JUL-1997; US-051919.
 PR 08-JUL-1997; US-051920.
 PR 08-JUL-1997; US-051925.
 PR 08-JUL-1997; US-051926.
 PR 08-JUL-1997; US-051928.
 DR WPI: 99-153780/13.

PR 08-JUL-1997; US-051929.
 PR 08-JUL-1997; US-051930.
 PR 08-JUL-1997; US-051931.
 PR 08-JUL-1997; US-051932.
 PR 08-JUL-1997; US-052732.
 PR 08-JUL-1997; US-052733.
 PR 08-JUL-1997; US-052793.
 PR 08-JUL-1997; US-052795.
 PR 08-JUL-1997; US-052803.
 PR 08-JUL-1997; US-052804.
 PR 18-AUG-1997; US-055722.
 PR 18-AUG-1997; US-055723.
 PR 18-AUG-1997; US-055947.
 PR 18-AUG-1997; US-055948.
 PR 18-AUG-1997; US-055949.
 PR 18-AUG-1997; US-055950.
 PR 18-AUG-1997; US-055953.
 PR 18-AUG-1997; US-055954.
 PR 18-AUG-1997; US-055954.
 PR 18-AUG-1997; US-055984.
 PR 18-AUG-1997; US-055984.
 PR 12-SEP-1997; US-058660.
 PR 12-SEP-1997; US-058661.
 PR 12-SEP-1997; US-058661.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Brewer LA, Ebner R, Fischer CL, Kyaw H, Lafleur DW, Li Y, Moore PA,
 PI Olsen HS, Rosen CA, Ruben SM, Shi Y, Soppet DR, Zeng Z;
 DR WPI: 99-120770/10.
 PT New isolated human genes and the secreted polypeptides they encode -
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders
 PT disclosure, Page 122; 464pp; English.
 CC This sequence represents a fragment of a secreted human protein encoded
 CC by the nucleic acid molecule detailed in the descriptor line. The gene
 CC can be used to generate fusion proteins by linking to the gene to a human
 CC immunoglobulin Fc portion (e.g. X27302) for increasing the stability of
 CC the fused protein as compared to the human protein only.
 CC The invention relates to 123 novel genes and their fragments (nucleic
 CC acid sequences: X27311-X27449) amino acid sequences X02650-Y02788) which
 CC are useful for preventing, treating or ameliorating medical conditions can be
 CC e.g. by protein or gene therapy. Also, pathological conditions can be
 CC diagnosed by determining the amount of the new polypeptides in a sample
 CC or by determining the presence of mutations in the new polynucleotides,
 CC Specific uses are described for each of the 123 polynucleotides, based on
 CC which tissues they are most highly expressed in (see X27311 for described
 CC uses).
 CC Sequence 53 AA;
 SQ
 Y02934 Length: 53 February 11, 2000 15:49 Type: P Check: 2639 ..
 1 TLVAGSPCSL SRWIMAFCH GELVOSDMES QEMERGCVL SHSLPCVY
 51 SPR
 IIA-SEQUENCE 1.0
 ID Y11970 standard; Protein: 113 AA.
 AC Y11970;
 DT 18-JUN-1999 (first entry)
 DE Human 5' EST secreted protein SEQ ID NO: 570.
 KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
 KW forensic; gene therapy; chromosome mapping; signal peptide; prostate;
 KW upstream regulatory sequence; cytokin activity; cell proliferation;
 KW differentiation; haematopoiesis regulation; tissue growth regulation;
 KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; anti-inflammatory; tumour inhibition.
 OS Homo sapiens.
 PN WO9906550-A2.
 PD 11-FEB-1999.
 PD 31-JUL-1998; IB1232.
 PR 01-AUG-1997; US-905144.
 PA (GEST) GENSET.
 PI Ducleert A, Dumas Milne Edwards J, Lacroix B;
 DR WPI: 99-153780/13.

DR N-PSDB; X40692.
PT New isolated prostate-derived nucleic acids - used to develop
PT products which may have cytokine, immune regulatory, haematopoiesis
PT regulating, anti-inflammatory or tumour inhibition activity
PS Claim 34; Page 660; 675pp; English.
CC X40438 to X40715 represent 5' expressed sequence tags (ESTs) for human
CC secreted proteins expressed in prostate, and encode the proteins given in
CC Y11716 to Y11993 respectively. The proteins given represent the signal
CC peptide and an N-terminal fragment of a secreted protein. The nucleic
CC acid sequences can be used for producing secreted human gene products.
CC They can also be used to develop products for diagnosis and therapy. The
CC proteins obtained may have cytokine activity, cell proliferation and
CC growth regulating activity, haematopoiesis regulating activity, tissue
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
CC receptor/ligand activity, anti-inflammatory activity, tumour inhibition
CC activity or other activities. The products can be used in forensic, gene
CC therapy and chromosome mapping procedures. The sequences can also be used
CC for obtaining corresponding promoter sequences. The nucleic acids
CC encoding the signal peptides can be used for directing extracellular
CC secretion of a polypeptide or the insertion of a polypeptide into a
CC membrane, or importing a polypeptide into a cell.
SO Sequence 113 AA;

Y11970 Length: 113 February 11, 2000 15:49 Type: P Check: 1817 ..

1 MANLFRKRV NPLLYLSRHT VKPRALSTYL FGSINGAPV AVEPGAAYS

51 LLSPLLPHL LPALGFKNKT VLKRRCKDCY LKRRGRMYV YCRTHPRHK

101 RHMXTLSIOS HAQ

!!AA-SEQUENCE 1.0
ID W96262 standard; Protein: 423 AA.
AC W96262;
DE 14-JUN-1999 (first entry)
DT Brn-3a polypeptide.
KW Brn-3a polypeptide.
KW CNS; central nervous system; parasympathetic nervous system;
KW PNS; central nervous system; parasympathetic nervous system;
KW development; injury; neurotrophic factor; nerve growth factor; NGF;
KW ciliary neurotrophic factor; CNF; brain-derived neurotrophic factor;
KW BDNF; neurotrophin; NT-3; NT-4; NT-5; neurodegenerative disease;
KW familial dysautonomia; infantile muscular dystrophy;
KW Parkinson's disease; Alzheimer's disease.
OS Homo sapiens.
PN W09905272-A1.
PD 04-FEB-1999;
PF 27-JUL-1998;
PR 10-DEC-1997; US-988476.
PR 25-JUL-1997; GB-015823.
PA (UNLO) UNIV COLLEGE LONDON.
PI Latchman DS; Smith MD;
DR WPI; 99-142828/12.
DR N-PSB; X09010.
PT New polypeptide comprising transcription factor Brn-3a, or its
PT derivative - useful for treating nervous system diseases, preventing
PT cellular apoptosis and increasing nerve regeneration following
PT neuronal damage
PS Disclosure; Page 61-62; 68pp; English.
CC Over expression of transcription factor Brn-3a can protect neuronal
CC cells from apoptosis. Brn-3a also specifically activates expression
CC of the Bcl-2 gene in neuronal cells and this activation is mediated
CC via a Brn-3a response element in the 5' regulatory region of the
CC Bcl-2 gene. Both the anti-apoptotic effect of Brn-3a and its ability
CC to activate expression of Bcl-2 are mediated by the N-terminal
CC domain of Brn-3a. Members of the Bcl-2 family perform critical roles
CC in the regulation of selective apoptosis during development of the
CC nervous system. The stimulation of Bcl-2 expression by Brn-3a in a
CC neuron specific manner and consequent protection of neuronal cells
CC from apoptosis suggests that Brn-3a may co-ordinate some aspects of
CC neuronal regeneration during development or following injury. The
CC elevation of Brn-3a expression by either pharmacological means
CC (compositions comprising one other therapeutic polypeptide e.g.

CC neurotrophic factors, nerve growth factor (NGF), ciliary neurotrophic
CC factor (CNTF), brain-derived neurotrophic factor (BDNF) and
CC neurotrophins NT-3 and NT-4/5) or gene therapy may represent a
CC method for treating human diseases associated with excessive
CC neuronal cell death and/or lack of nerve regeneration, especially
CC neurodegenerative diseases such as familial dysautonomia and
CC infantile muscular dystrophy, and Parkinson's and Alzheimer's
CC disease.
SO Sequence 423 AA;

W96262 Length: 423 February 11, 2000 15:49 Type: P Check: 3324 ..

1 MMSNSKOPH FAPHPTLPEH KYPSLSHSSSE AIRRACLETP PLOSNFASL

51 DETLLARAFA LAVDIAVSQ GKSHPEKPPA TYHTMNSVPC TSTSTVPLRH

101 HHHHHHHHQA LEPDLDLHDI SSPLALIMAG AGAGAGAGAA AGGGGAHDGP

151 GGGGGPPGGG GPGGGPPGG GGGGPGGGG GPGGGLLGS AHPHMHSL

201 GHLSPHAAA AMNPSGLPH PGLVAAAHH GAAAAAAA AGVAAASAA

251 AAVGAAGIA SICSDTDPR ELEAFERPK QRIKLGVTQ ADVGSALANL

301 KIPGVSLSQ STICRESLT LSHNMVIALK PILQAMLEEA EGAKREKMK

351 PELFNGGEKK KKRSTIAPE KSLSEAFYAV QPRSEKIA AIAEKDLK

401 NVVRVFCNQ ROKKRWKFS ATY

!!AA-SEQUENCE 1.0
ID Y05204 standard; peptide: 24 AA.
AC Y05204;
DE 17-JUN-1999 (first entry)
DT Human CLIP immunomodulatory peptide.
KW CLIP; immunomodulatory peptide; immune disorder; TH1 cell reduction;
KW MHC class II-associated invariant chain protein; transplant rejection;
KW TH2 cell; immune response; T cell response; T cell activation; arthritis;
KW autoimmune disorder; autoimmune diabetes; multiple sclerosis;
KW systemic lupus erythematosus; cytokine production; human.
OS Homo sapiens.
PN CA2205680-A.
PD 16-NOV-1998.
PF 16-MAY-1997; CA-205680.
PR (UYWO-) UNIV WESTERN ONTARIO.
PA Singh B;
PI WPI; 99-244657/21.
DR Use of an MHC class II-associated invariant chain protein (CLIP) for
DR treating or preventing an immune disorder
PT Claim 3; Fig 1; 62pp; English.
PS This sequence represents the human CLIP immunomodulatory peptide.
CC The invention relates to a method of treating or preventing an immune
CC disorder comprising administering an MHC class II-associated invariant
CC chain protein (CLIP) to an affected/susceptible mammal. The CLIP peptides
CC are useful as immunogenic compositions for: (1) increasing the level of
CC TH2 cells; (2) decreasing the level of TH1 cells; (3) preventing or
CC reducing an immune response; (4) reducing or preventing a T cell response
CC to an antigen; (5) reducing or preventing antigen presentation by APCs;
CC and (6) treating or preventing a disorder associated with a T cell
CC response. In mammals, preferably humans. The immune disorders involve
CC autoimmune disorders, including autoimmune diabetes, arthritis, multiple
CC sclerosis and systemic lupus erythematosus. The CLIP peptides are useful
CC for screening for activity of CLIP fragments/analogues, and for assaying
CC for ability to affect cytokine production such as interleukin-2 or
CC interferon gamma.
SO Sequence 24 AA;

Y05204 Length: 24 February 11, 2000 15:49 Type: P Check: 3125 ..

1 KRPKPVSK MRMAPPLMQ ALPM

11AA SEQUENCE 1.0
 ID Y11993 standard; Protein: 102 AA.
 AC Y11993;
 DT 18-JUN-1999 (first entry)
 DE Human 5' EST secreted protein SEQ ID NO: 593.
 KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
 KW forensic; gene therapy; chromosome mapping; signal peptide; prostate;
 KW upstream regulatory sequence; cytokine activity; cell proliferation;
 KW differentiation; haematopoiesis regulation; tissue growth regulation;
 KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; anti-inflammatory; tumour inhibition.
 OS Homo sapiens.
 PN WO9906350-A2.
 PD 11-FEB-1999.
 PR 31-JUL-1998; 1B1232.
 PR 01-AUG-1997; US-903144.
 PA (GENSET) GENSET. Milne Edwards J, Lacroix B;
 PI Nucleot A, Dumas Milne Edwards J, Lacroix B;
 DR N-PSDB: X40715.
 DR WPI: 99-153780/13.
 PT New isolated prostate-derived nucleic acids - used to develop
 PT products which may have cytokine, immune regulatory haematopoiesis
 PT regulating, anti-inflammatory or tumour inhibition activity
 PS Claim 34: Page 675; 675pp: English.
 CC X40438 to X40715 represent 5' expressed sequence tags (ESTs) for human
 CC secreted proteins expressed in prostate, and encode the proteins given in
 CC Y11716 to Y11993 respectively. The proteins given represent the signal
 CC peptide and an N-terminal fragment of a secreted protein. The nucleic
 CC acid sequences can be used for producing secreted human gene products.
 CC They can also be used to develop products for diagnosis and therapy. The
 CC proteins obtained may have cytokine activity, cell proliferation and
 CC differentiation activity, haematopoiesis regulating activity, tissue
 CC growth regulating activity, reproductive hormone regulating activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
 CC receptor/ligand activity, anti-inflammatory activity, tumour inhibition
 CC activity or other activities. The products can be used in forensic, gene
 CC therapy and chromosome mapping procedures. The sequences can also be used
 CC for obtaining corresponding promoter sequences. The nucleic acids
 CC encoding the signal peptides can be used for directing extracellular
 CC membrane, or importing a polypeptide into a cell.
 CC Sequence 102 AA.
 Y11993 Length: 102 February 11, 2000 15:49 Type: P Check: 4119 ..

1 MDPFESVTE NFEYDILAE CYIGDIYVG TVFISITYSV IFALIGVNL
 51 LVYFALTNSK KPKSVTDIYL LNLASDLIF VAIPEFMTYH LINEKGLHVA
 101 MC
 11AA SEQUENCE 1.0
 ID Y05545 standard; Protein: 243 AA.
 AC Y05545;
 DT 05-JUL-1999 (first entry)
 DE Wheat type III glutathione transferase subunit ICP/ICV.
 KW Glutathione transferase; GST; glutathione peroxidase; wheat; ICP;
 KW ICP; herbicide resistance; transgenic plant.
 OS Triticum aestivum.
 PN MO9914337-A2.
 PD 25-MAR-1999.
 PR 16-SEP-1997; GB-019727.
 PR 16-SEP-1997; GB-019727.
 PA (RHON) RHONF-POULENC AGRIC LTD.
 PI Cole DJ Cummings I, Edwards R;
 DR N-PSDB: X25152.
 DR N-PSDB: X25152.
 PT New isolated glutathione transferase subunit polynucleotides
 PT Claim 3, Page 100, 101pp: English.
 CC The present sequence represents ICP and ICV, wheat glutathione
 CC transferase (GST) subunits that resemble the type III GSTs of maize.
 CC ICP/ICV cDNA (see X25152) was isolated from a cDNA library prepared
 CC from fenchlorazole-ethyl (herbicide safener) treated wheat shoots.

CC The invention provides wheat GST subunits (see Y05537-45) active in
 CC herbicide metabolism. This is fundamental to understanding GST
 CC detoxification in wheat and in the development of transgenic
 CC herbicide resistant plants expressing wheat GSTs. The invention
 CC provides methods of identifying compounds capable of metabolism by
 CC GST, or compounds that induce GST expression in gramaceous plants,
 CC and for determining the GST level in a sample of seed or flour.
 CC Transgenic plants, host cells used for production of GST subunits
 CC and GST dimeric proteins, and vectors are also provided.
 CC Sequence 243 AA;
 Y05545 Length: 243 February 11, 2000 15:49 Type: P Check: 5919 ..

1 MAGGELTKL GWNAPGVSPY VLRAQMALAV KGLSTYDUE DWSSTDLIL
 51 ASNPYKATP VLIINGRPVC ESLLILEYLD DAVGLANGK PILPADPYSR
 101 AVAREMAAV NDKLPSCGT ILKTKOER AGKEETLSG LRLBAVAAE
 151 CSSEGEAPFP FGDAIGFD IALGCLPWF ENAGRLAGLG PILDAPTPK
 201 LAAMERFSV AEPKALLPG VDKLEIYIT ALYKMNIAV TGN
 11AA SEQUENCE 1.0
 ID Y07041 standard; Protein: 207 AA.
 AC Y07041;
 DT 02-JUL-1999 (first entry)
 DE Breast cancer associated antigen precursor sequence.
 KW Breast cancer associated antigen; diagnosis; research; treatment; human;
 KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
 KW prostate cancer.
 OS Homo sapiens.
 PN WO9904265-A2.
 PD 28-JAN-1999.
 PR 15-JUL-1998; US-014679.
 PR 22-JUN-1998; US-102322.
 PR 17-JUL-1997; US-896164.
 PR 10-OCT-1997; US-061599.
 PR 10-OCT-1997; US-061765.
 PR 10-OCT-1997; US-948705.
 PR 11-OCT-1997; GB-021697.
 PA (LUDM-) LUDMIG INST CANCER RES.
 PI Chen Y, Gout I, Gure A, Ohare M, Obara Y, Old LJ,
 PI Pfreundschuh M, Sahlin U, Scanlan MJ, Stockert E,
 PI Tureci O;
 DR WPI: 99-132448/11.
 PT New isolated cancer associated nucleic acids and polypeptides -
 PT isolated using sera from cancer patients, used to develop products
 PT for the diagnosis, monitoring or treatment of cancers
 PS Disclosure: Page 426-427; 78pp: English.
 CC The invention relates to a method for diagnosing a disorder characterised
 CC by expression of a human cancer associated antigen precursor coded for by
 CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
 CC biological sample isolated from a subject with an agent that specifically
 CC binds to the NAM, an expression product or a fragment of an expression
 CC product complexed with an HLA molecule; and (b) determining the
 CC interaction between the agent and the NAM or the expression product as a
 CC determination of the disorder. The products and methods can be used in
 CC the diagnosis, monitoring, research, or treatment of conditions
 CC characterised by the expression of various cancer associated antigens.
 CC The invention provides nucleic acid sequences and encoded polypeptides
 CC which are cancer associated antigen precursors expressed in human breast
 CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
 CC lung cancer.
 CC Sequence 207 AA;
 Y07041 Length: 207 February 11, 2000 15:49 Type: P Check: 1679 ..

1 MSPILRLL ALALQAPAQ AVSPDADG HQRKRVWID VYTRATQPR
 51 EYVPLTEL MGTAKOLVP SCVTYORCGG CCPDGEVCY PTGOHVRMO
 101 ILMTRPSSQ LQMSLEBHS QCECRKRRKD SAVKPRRAAT PHHRPQPRSV

151 PGMSARGAP SPADITHPTP APGPSANMAP STTSALTGPB AAAADAAAS
201 SVAKGA
11AA_SEQUENCE 1.0
ID Y07007 standard; Protein: 177 AA.
AC Y07007;
DE Breast cancer associated antigen precursor sequence.
KW Breast cancer; colon cancer; gastric cancer; lung cancer;
OS Homo sapiens.
PN WO9904265-A2.
PR 15-JUL-1998: U14679.
PR 22-JUN-1998: US-102322.
PR 17-JUL-1997: US-896164.
PR 10-OCT-1987: US-061599.
PR 10-OCT-1997: US-061765.
PR 10-OCT-1997: US-948705.
PR 11-OCT-1997: GB-021697.
PA (JUDW-) LUDWIG INST CANCER RES.
PI Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ,
PI Pfreundschuh M, Sahin U, Scanlan MJ, Stockert E,
PI Tureci O;
DR WPI: 99-132448/J1.
PT New isolated cancer associated nucleic acids and polypeptides -
PT isolated using sera from cancer patients, used to develop products
PT for the diagnosis, monitoring or treatment of cancers
PS Disclosure; Page 391-392, 787pp; English
CC The invention relates to a method for diagnosing a disorder characterised
CC by expression of a human cancer associated antigen precursor coded for by
CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC biological sample isolated from a subject with an agent that specifically
CC binds to the NAM, an expression product or a fragment of an expression
CC product complexed with an HLA molecule; and (b) determining the
CC interaction between the agent and the NAM or the expression product as a
CC determination of the disorder. The products and methods can be used in
CC the diagnosis, monitoring, research, or treatment of conditions
CC characterised by the expression of various cancer associated antigens.
CC The invention provides nucleic acid sequences and encoded polypeptides
CC which are cancer associated antigen precursors expressed in human breast
CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC lung cancer.
SQ Sequence 177 AA;
Y07007 Length: 177 February 11, 2000 15:49 Type: P Check: 4692
1 LCCHMYCKS CWNEYLITRI EQNLVLCNCT PIADCPAOT GAFIRAIYVS
51 PEVISKYKAL LRGVYSCSN LTMCTNPOGC DRILCRQGLG CGTSCSKCGW
101 ASCNCFEPE AHYPASCGHM SQWDDGGYV DGMVSAKHL AKLISKRCPS
151 CQPIENEGC LHMTCACNKH GFCNRCL
11AA_SEQUENCE 1.0
ID W93953 standard; Protein: 352 AA.
AC W93953;
DE Human regulatory molecule HRM-9 protein.
KW Human regulatory molecule; HRM-9; cytostatic activity; immune modulator;
KW transcription factor; enhancer; cell proliferation stimulation; cancer;
KW leukemia; lymphoma; myeloma; adenocarcinoma; sarcoma; bladder; bone;
KW brain; lung; liver; ovary; skin; teratocarcinoma; immune response;
KW allergy; asthma; diabetes; multiple sclerosis; Grave's disease;
OS Homo sapiens.
PN WO991658-A2.
PR 01-APR-1999.

PF 22-SEP-1998: U19839.
PR 23-SEP-1997: US-933750.
PA (INCY-) INCYTE PHARM INC.
PI Au-Young J, Bandman O, Guegler KJ, Hillman JL, Lal P,
PI Shah P, Yue H;
DR WPI: 99-254710/21.
DR N-PSDB: X24067.
PS New human regulatory molecules
PT Claim 1; Page 75; 76pp; English.
CC This invention describes novel human regulatory molecules (HRM) which
CC have cytostatic activity and act as immune modulators, transcription
CC factors or enhancers. The HRMs can be used to stimulate cell
CC proliferation. Antagonists and agonists of the proteins of the invention
CC can be used to treat cancer. The encoding nucleic acids can be used in
CC microarrays to detect polynucleotides (and their expression levels) that
CC encode HRMs in a biological sample. The HRMs and microarrays can be used
CC to diagnose, treat or prevent cell proliferation diseases especially cancer,
CC e.g. leukemia, lymphoma, myeloma, adenocarcinoma, sarcoma, cancer of e.g.
CC bladder, bone, brain, lung, liver, ovary, skin, etc. teratocarcinoma, or
CC to treat or prevent immune responses e.g. allergies, asthma, diabetes,
CC multiple sclerosis, Grave's disease or myasthenia gravis.
SQ Sequence 352 AA;
W93953 Length: 352 February 11, 2000 15:49 Type: P Check: 7759
1 MHVAPASUR LGTGTLNPPS PTCITRLALP PAAPSLAM SOSRRAEAP
51 PLEBEDSGTF SLGKMITAKP GKTPIQVLHE YGMKTKNIPV YECERSDVOI
101 HVTFFFRVT VGDITCGEG TSKKLAKHRA AEAANILKA NASICPAVVD
151 PLMPDSKRP KQNLNPIGSL QELAIHGW RLEPTIOSEG GPARKREYTT
201 ICHLEFMT GKASKKOAK RNAAEFLAK FSNISPEHNI SLTNVGHSL
251 GCTWHSLRNS PGEKINILKR SLISIPNTDY IQLISEIAKE OGENTYLDI
301 DELSANGYOQ CLABLSTSPI TVCHSGSISG GNAOSDAHN ALOYLKIIE
351 RK
11AA_SEQUENCE 1.0
ID Y04959 standard; Protein: 273 AA.
AC Y04959;
DE 06-JUL-1999 (first entry)
DE Mycobacterium species Protein sequence 42D.
KW Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
KW hybridisation; detection; vaccine; immunisation; infection.
OS Mycobacterium sp.
PN WO9909186-A2.
PR 25-FEB-1999.
PR 14-AUG-1998: F01813.
PR 11-SEP-1997: FR-011325.
PR 14-AUG-1997: FR-010404.
PA (INSP) INST PASTEUR.
PI Gicquel B, Lhm EM, Pelletier V, Portnoi D, Goguet de la Salmoniere Y,
PI Guigneno B;
DR WPI: 99-181045/15.
DR N-PSDB: X34210.
PS Mycobacterial DNA vectors containing reporter constructs - for
PT identifying coding or promoter sequences involved in
PT infection-associated protein expression
PS Claim 3; Fig 42D; 309pp; French.
CC Sequences Y04742-Y05000 and Y07201-Y07204 represent secreted proteins
CC from various Mycobacterium species microorganisms. The encoding
CC nucleotide sequences can be used as primers and probes for methods
CC for detecting and identifying mycobacteria, especially belonging to
CC the M. tuberculosis complex. The encoded proteins can be used in
CC vaccines for immunisation against a bacterial or viral infection.
SQ Sequence 273 AA;
Y04959 Length: 273 February 11, 2000 15:49 Type: P Check: 8454


```

1 MANKRAGQ PLEPSRDD HMOGHMILAR LGRVLRPGG VELTRILLAR
51 AEVTDADVLE LAPGLGRFA EILARNRSTY VGSADSPNNA NLVHVHLAGR
101 GGVRTDADP TGLSDASADV VIGEMALTMQ GNAKHTIYA EAAVLRPGG
151 RYAIHELALV PDDVAEQVRT DLROSLARAL KVAARPLIVA EMSHLLAGHG
201 LVEHVHTAS MALLQPRVYI ADEGLLGLAR FAGNLIHRA ARRVLLMRH
251 TFRHREHRT AVAIVAKPH VDS

!!AA_SEQUENCE 1.0
ID Y04960 standard; Protein: 280 AA.
AC Y04960;
DE Mycobacterium species protein sequence 42F.
KW Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
OS hybridisation; detection; vaccine; immunisation; infection.
PN WO9909186-A2.
PD 25-FEB-1999.
PF 14-AUG-1998; F01813.
PR 11-SEP-1997; FR-011325.
PA (INSP) INST PASTEREUR.
PI Gicquel B, Lhm EM, Pelicic V, Portnoi D, Goguet de la Salmoniere Y,
DR Guigueno A;
DR WPI; 99-181045/15.
DR N-PSDB; X34211.
PT Mycobacterial DNA vectors containing reporter constructs - for
PT identifying coding or promoter sequences involved in
PS infection-associated protein expression
PS Claim 32; Fig 42F; 309pp; French.
CC Sequences Y04742-Y05000 and Y07201-Y07204 represent secreted proteins
CC from various Mycobacterium species microorganisms. The encoding
CC nucleotide sequences can be used as primers and probes for methods
CC for detecting and identifying mycobacteria, especially belonging to
CC the M. tuberculosis complex. The encoded proteins can be used in
CC vaccines for immunisation against a bacterial or viral infection.
SQ Sequence 280 AA;

Y04960 Length: 280 February 11, 2000 15:49 Type: P Check: 8320
1 KPAVGSMAK KRGNAQPIR LSPRDDHMO GHMILARLGR RVLRPGVEL
51 TRILARAEV TDADVLEAP GLGRFAEIL ARNRSTYGA ESPDNANILV
101 RYVLAGRGDV RTDADDTGL SDAADVIG EAMTMOGNA AKHTIYAENA
151 RVLRPGGRYA IHELALVPDD VAEQVRTDL QSLARALKVN ARELVAEMS
201 HILAGHGLV EHVHTASML LQPRVYIDE GLGLARFAG NLIHRAAR
251 RYVLRHRTFR RHRELTAVA IVAHPRVDS

!!AA_SEQUENCE 1.0
ID Y04901 standard; Protein: 72 AA.
AC Y04901;
DE Mycobacterium species protein sequence 31C.
KW Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
OS hybridisation; detection; vaccine; immunisation; infection.
PN WO9909186-A2.
PD 25-FEB-1999.
PF 14-AUG-1998; F01813.
PR 11-SEP-1997; FR-011325.
PA (INSP) INST PASTEREUR.
PI Gicquel B, Lhm EM, Pelicic V, Portnoi D, Goguet de la Salmoniere Y,
DR Guigueno A;
DR WPI; 99-181045/15.

```

```

DR N-PSDB; X34154.
PT Mycobacterial DNA vectors containing reporter constructs - for
PT identifying coding or promoter sequences involved in
PS infection-associated protein expression
PS Claim 32; Fig 31C; 309pp; French.
CC Sequences Y04742-Y05000 and Y07201-Y07204 represent secreted proteins
CC from various Mycobacterium species microorganisms. The encoding
CC nucleotide sequences can be used as primers and probes for methods
CC for detecting and identifying mycobacteria, especially belonging to
CC the M. tuberculosis complex. The encoded proteins can be used in
CC vaccines for immunisation against a bacterial or viral infection.
SQ Sequence 72 AA;

Y04901 Length: 72 February 11, 2000 15:49 Type: P Check: 6286
1 ANTREGSATA ALPAPMLNN TSOPRTDFVC GRINDPNSTN TRIATKALA
51 PGISIRATRR SAASISARRA SR

!!AA_SEQUENCE 1.0
ID Y04743 standard; Protein: 110 AA.
AC Y04743;
DE Mycobacterium species protein sequence 142.
KW Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
OS hybridisation; detection; vaccine; immunisation; infection.
PN WO9909186-A2.
PD 25-FEB-1999.
PF 14-AUG-1998; F01813.
PR 11-SEP-1997; FR-011325.
PA (INSP) INST PASTEREUR.
PI Gicquel B, Lhm EM, Pelicic V, Portnoi D, Goguet de la Salmoniere Y,
DR Guigueno A;
DR WPI; 99-181045/15.
DR N-PSDB; X34001.
PT Mycobacterial DNA vectors containing reporter constructs - for
PT identifying coding or promoter sequences involved in
PS infection-associated protein expression
PS Claim 32; Fig 1; 309pp; French.
CC Sequences Y04742-Y05000 and Y07201-Y07204 represent secreted proteins
CC from various Mycobacterium species microorganisms. The encoding
CC nucleotide sequences can be used as primers and probes for methods
CC for detecting and identifying mycobacteria, especially belonging to
CC the M. tuberculosis complex. The encoded proteins can be used in
CC vaccines for immunisation against a bacterial or viral infection.
SQ Sequence 110 AA;

Y04743 Length: 110 February 11, 2000 15:49 Type: P Check: 5429
1 MAGQEELELR FVPLYTLAE ASRTLYVPR TLTADNGE RRPANNAVQ
51 GQPIAFDAYS VAOLFQDVTG ARVAGVQPR HIRRVRLRG PLGVCGLRH
101 PRQFAGTISO

!!AA_SEQUENCE 1.0
ID Y07239 standard; Protein: 141 AA.
AC Y07239;
DE Mycobacterium species protein sequence 31C.
KW Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
OS hybridisation; detection; vaccine; immunisation; infection.
PN WO9909186-A2.
PD 25-FEB-1999.
PF 14-AUG-1998; F01813.
PR 11-SEP-1997; FR-011325.
PA (INSP) INST PASTEREUR.
PI Gicquel B, Lhm EM, Pelicic V, Portnoi D, Goguet de la Salmoniere Y,
DR Guigueno A;
DR WPI; 99-181045/15.

```

DR WPI: 99-207107/18.
DR N-PSDB: X29975.
PT New allelic variant of human STAT3 useful in treating autoimmune or
inflammatory diseases
PS Disclosure: Page 8-9; 22pp; English.
CC This sequence represents a fragment of the placental isoform of human
Signal Transducer and Activator of Transcription (STAT3) intracellular
transcription factor (Akira et al., Cell 77, 63-71 (1994)). The
invention relates to isolation of allelic variants of the placental
hSTAT3 sequence. hSTAT3 plays a role in the upregulation of hepatic
acute-phase proteins, growth arrest of monocytic cells and in the
survival of myeloma cells and so may be used to treat or diagnose
autoimmune or inflammatory diseases.
SQ Sequence 141 AA;

Y07239 Length: 141 February 11, 2000 15:49 Type: P Check: 2144

1 WLDNIIDLVK KYIALMNEG YIMGFISKER ERAILSTRPP GTFLRFSES
51 SKEGVTFTW VKDISGKTQ IOSVEPYTKQ QLNMSFAEI IMGYKIMDAT
101 NILVSPLYVL YPDIPKEAF GKICRPESOE HPEADPSAA P

11AA-SEQUENCE 1.0
ID Y07240 standard; Protein; 141 AA.

AC Y07240: 1999 (first entry)
DE Fragment of human hepatic STAT3 protein.
KW Placenta; isoform; human; STAT3; Intracellular; transcription factor;
KW Signal Transducer and Activator of Transcription; allele; growth arrest;
KW hepatic acute-phase protein; monocytic cell; myeloma; autoimmune disease;
KW inflammation.
OS Homo sapiens.
PN EP-906953-A1.
PD 07-APR-1999.
PF 16-SEP-1997; 116061.
PR 16-SEP-1997; EP-116061.
PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
PI Della Pietra L, Serlupi-Crescenzi O;
DR WPI: 99-207107/18.
DR N-PSDB: X29976.
PT New allelic variant of human STAT3 useful in treating autoimmune or
inflammatory diseases
PS Claim 2: Page 10-11; 22pp; English.
CC This sequence represents a fragment of the hepatic allelic isoform
of human Signal Transducer and Activator of Transcription (STAT3)
intracellular transcription factor (Akira et al., Cell 77, 63-71
(1994)). The invention relates to isolation of allelic variants of
the placental hSTAT3 sequence. hSTAT3 plays a role in the upregulation
of hepatic acute-phase proteins, growth arrest of monocytic cells and
in the survival of myeloma cells and so may be used to treat or diagnose
autoimmune or inflammatory diseases.
SQ Sequence 141 AA;

Y07240 Length: 141 February 11, 2000 15:49 Type: P Check: 2614

1 WLDNIIDLVK KYIALMNEG YIMGFISKER ERAILSTRPP GTFLRFSES
51 SKEGVTFTW VKDISGKTQ IOSVEPYTKQ QLNMSFAEI IMGYKIMDAT
101 NILVSPLYVL YPDIPKEAF GKICRPESOE HPEADPSAA P

11AA-SEQUENCE 1.0
ID Y07241 standard; Protein; 141 AA.
AC Y07241:
DE 06-JUL-1999 (first entry)
DE Fragment of mouse hepatic STAT3 protein.
KW Placenta; isoform; human; STAT3; Intracellular; transcription factor;
KW Signal Transducer and Activator of Transcription; allele; growth arrest;
KW hepatic acute-phase protein; monocytic cell; myeloma; autoimmune disease;
KW inflammation.
OS Homo sapiens.
PN EP-106953-A1.

PD 07-APR-1999.
PF 16-SEP-1997; 116061.
PR 16-SEP-1997; EP-116061.
PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
PI Della Pietra L, Serlupi-Crescenzi O;
DR WPI: 99-207107/18.
DR N-PSDB: X29977.
PT New allelic variant of human STAT3 useful in treating autoimmune or
inflammatory diseases
PS Disclosure: Page 12-13; 22pp; English.
CC This sequence represents a fragment of the hepatic allelic isoform
of mouse Signal Transducer and Activator of Transcription (STAT3)
intracellular transcription factor (Akira et al., Cell 77, 63-71
(1994)). The invention relates to isolation of allelic variants of
the placental human STAT3 sequence. hSTAT3 plays a role in the
upregulation of hepatic acute-phase proteins, growth arrest of monocytic
cells and in the survival of myeloma cells and so may be used to treat
or diagnose autoimmune or inflammatory diseases.
SQ Sequence 141 AA;

Y07241 Length: 141 February 11, 2000 15:49 Type: P Check: 3064

1 WLDNIIDLVK KYIALMNEG YIMGFISKER ERAILSTRPP GTFLRFSES
51 SKEGVTFTW VKDISGKTQ IOSVEPYTKQ QLNMSFAEI IMGYKIMDAT
101 NILVSPLYVL YPDIPKEAF GKICRPESOE HPEADPSAA P